2017

# Using rational expectations storage model to explain natural gas price

Shufen Chen
*Iowa State University*

**Using rational expectations storage model to explain natural gas price**

by

**Shufen Chen**

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Bruce A. Babcock, Major Professor
Sebastien Pouliot, Major Professor
Chad E. Hart
Dermot J. Hayes
Gabriel E. Lade

Iowa State University

Ames, Iowa

2017

TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEGEMENTS

First and foremost, I would like to express my sincere gratitude to my major professor, Professor Bruce A. Babcock. He is the best major professor I could ask for. His valuable insights and comments are critical for my dissertation. I couldn't have made it without his continuous encouragement and support. I would like to thank Dr. Sebastien Pouliot for being so supportive. Together with Professor Babcock, Dr. Pouliot helped edit my dissertation draft and provides many valuable suggestions. I am grateful to Dr. Dermot J. Hayes, Dr. Chad E. Hart, Dr. Gabriel E. Lade for their helpful comments and constructive criticisms regarding this study.

I would like to thank my husband, Yue Zhao, for his understanding, trust, continuous support and encouragement throughout the study.

# ABSTRACT

Natural gas is a key energy source for residential, commercial, electric power and industrial use. Residential, commercial and electric power sector consumption is primarily driven by weather conditions and displays obvious seasonal patterns while production is relatively stable throughout the year. As weather condition is uncertain and both consumption and production are inelastic in the short term, natural gas price is quite volatile especially in the heating season. Due to an imbalance between production and consumption, storage plays an important role in ensuring availability and smoothing price between low and peak consumption seasons. Storage is also a key driver of price volatility. The existing literature confirms the importance of inventory and weather conditions in determining price and its variance. Most studies to date use time series models and focus on the historical price realizations while providing little insight into how price patterns are determined by market participants' behavior. In addition, the impact of inventory and weather variables on price volatility has not been analyzed in detail.

This thesis aims to construct a model that can mimic the major market participants' behavior and reproduce the natural gas price with mean and standard deviation patterns consistent with historical observations: higher average level and standard deviation in peak consumption season. We construct a monthly rational-expectations competitive storage model to better reflect monthly variations in price. Natural gas consumption and production are specified in a way that the current period volume is highly correlated with previous period volume so as to capture the stickiness and gradual change in natural gas markets. Imposing non-arbitrage condition, price is inter-temporally correlated. Net storage cost consists of both physical storage

cost and convenience yield obtained from holding stock at hand. Normal storage level for each month is introduced to reflect the yearly cycling of natural gas inventory and is used in the convenience yield calibration. It denotes the normal storage level each month that is needed to balance seasonal demand-supply relationship. Convenience yield is high if the inventory falls below normal storage level and high convenience yield pushes up the price and decreases current consumption to accumulate more natural gas for future use.

The model is solved using numerical methods because analytical solutions are not feasible. In order to validate the result, accuracy tests are conducted and the major assumptions are tested as well. The model's approximation errors are reasonable. The model is further validated by comparing simulated price series with historical observations by using historical weather variables in the solved model. The simulated model generates prices that largely replicate the key features of historical data, including the price level, price variance, price sensitivity under unusual weather conditions and price autocorrelation.

Weather conditions and total natural gas availability are the main drivers for price and price standard deviation. The model finds that in winter high heating degree days (HDD) or low inventory drives price and price volatility higher while price and its variance decrease with low HDD and high inventory. The case is similar in summer with cooling degree days (CDD) instead of HDD as the weather variable. When inventory is low, weather shocks have a larger impact on price than when inventory is high. The effect is more pronounced in winter than in summer because the supply is tighter in heating season.

Using the validated competitive storage model, this thesis further assesses the potential impact of LNG export on the U.S. domestic natural gas market. Given the large pricing spread between the United States and the rest of world, along with policy promotion and the completion of LNG facilities, U.S. LNG exports are poised to expand dramatically. This study covers two major types of LNG export scenarios: exogenous fixed volume and endogenous export volume depending on the price spread between US and world prices. Four export scenarios are analyzed and compared with the benchmark scenario of no LNG export. The first two scenarios are fixed export volume with 6 bcf/day and 12 bcf/day respectively, to be consistent and comparable with an EIA 2014 report and the existing literature. One of the endogenous scenario scenarios assumes no consumption and supply growth for importing countries and the other one assume 100% increase of demand and 50% increase of supply in LNG importing countries by 2036.

Because of high shipping cost and inelastic natural gas demand in importing countries, the U.S. LNG is not competitive under current market condition, if no growth is expected. The U.S. LNG export volume is very small and decreases over time. Due to small export volumes, the domestic price impact is minimal. For all scenarios analyzed in this study, the long-term price impact is less than 8%, or around $0.33 per thousand cubic feet. In the long-term, the endogenous export with growth assumption scenario shows the largest price increase compared to the no export benchmark scenario. The export level is around 12 bcf per day.

The U.S. domestic price variance becomes smaller if an endogenous export sector is added while the price variance becomes higher under fixed export volume scenarios. If the LNG export is endogenously determined, when domestic price increased, LNG export decreases. This provides an additional buffer to the U.S. domestic market if there is shock to push up natural gas consumption and price. In contrast, fixed volume export makes the total natural gas consumption less responsive to price change and thus increases price variance.

Most of the LNG export volumes will be satisfied by production increases instead of domestic consumption reductions in the long term. In all four scenarios analyzed in this study, production catches up gradually in response to price increase due to LNG export. In the beginning period when production is constrained by production capacity, most of the export is covered by domestic consumption reduction. In the long term, as production increase, domestic consumption recovers to similar level as in the no export scenario.

# CHAPTER 1. INTRODUCTION

Natural gas is an important energy source for the residential, commercial, electric power and industrial sectors of the U.S. economy. The natural gas market has garnered a lot of attention recently due to its growing importance as an energy source and the many significant changes the market has gone through in the past decade. With the technology advance in hydraulic fracturing and horizontal drilling, domestic production of natural gas has increased dramatically since 2005. Industrial and electrical power sector natural gas usage increased in response to significant price decreases. Natural gas price is one of the most volatile among energies due to inelastic supply and highly weather-driven consumption in the short period. The aim of this research is to better understand the dynamics of the natural gas market and associated price behavior.

## 1.1 Natural Gas Market Fundamentals

Weather has substantial impact on natural gas storage and prices. On the consumption side, residential and commercial uses of natural gas are directly affected by temperature. When temperature is high, cooling consumption is large while low temperatures increase heating needs of residential and commercial users. Natural gas is also an important feedstock for electricity generation and is replacing the leading role of coal because of low gas prices and environmental issues involved with coal. Electricity consumption has strong seasonality and closely related to weather conditions. On the supply side, within a short period, natural gas supply is usually quite stable. But over time supply has changed dramatically. Since 2005, with the shale gas revolution, U.S. natural gas supply has increased by about 50%, from 19 trillion cubic feet (tcf) in 2005 to 28 tcf in 2016. This supply is affected by random shocks like extreme weather. For example, Hurricane Katrina in 2005 August and Hurricane Rita in 2009 greatly decreased the short-term natural gas supply as some of the Gulf Coast natural gas infrastructure was destroyed. Storage is important in balancing supply and consumption and smoothing natural gas prices fluctuations across different periods. In winter, natural gas consumption is high and storage is used to satisfy seasonal peak demands. In summer, spring and fall, natural gas is stored for winter use. Storage accounts for about 40% of winter month natural gas supply.

According to U.S. Energy Information Administration (EIA), natural gas delivered to consumers can be categorized as commercial consumption, residential consumption, electric power consumption, industrial consumption and vehicle fuel consumption. A small proportion of the marketed natural gas is used for plant processing and distribution. We focus on natural gas end-use in this study.

The residential sector is large and exhibits quite variable consumption patterns. All private dwellings' usage of natural gas is categorized as residential consumption, including house or apartment heating, cooling, cooking, and all other household uses. The residential sector accounted for 17.4% of the total use in 2016. Residential natural gas consumption has an obvious seasonal pattern. In winter, heating consumption drives natural gas consumption high. Commercial natural gas consumption is defined as the "gas used by nonmanufacturing establishments or agencies primarily engaged in the sale of goods or services. Included are such establishments as hotels, restaurants, wholesale and retail stores and other service enterprises; gas used by local, State, and Federal agencies engaged in nonmanufacturing activities" (EIA, 2016). Vehicle fuel natural gas consumption is included in commercial gas consumption through 1996 by EIA. Vehicle fuel natural gas consumption accounted for 1.2% of the total commercial consumption and 0.2% of total gas consumption in 2016. To better use the data and because vehicle fuel only accounts for a very small proportion, this study treats vehicle fuel natural gas consumption as part of commercial consumption. Gas consumption in commercial sector follows the same pattern as residential sector and is mainly driven by weather factors. In winter, heating consumption in the commercial sector drives up the commercial natural gas consumption. In 2016, commercial sector consumed 12.6% of the total natural gas delivered to consumers.

Electric power sector is the largest sector, accounting for 39.5% of the total natural gas consumption in 2016. Gas consumed in this sector is categorized in two types: electricity-only plants and combined heat and power plants. For combined heat and power plants, if they are identified as primarily for residential or commercial use, natural gas use will be categorized as residential or commercial use instead (EIA, 2016). When the price of natural gas was high relative to coal, natural gas was used as a peak load feedstock for electric power. Since the significant price reduction due to the shale gas boom, some efficient gas-burning combined cycle plants are used as base load generation (EIA, 2012). We can see a clear upward trend in natural gas consumption in the power sector in the recent decade. Natural gas consumption in the power sector is weather responsive and has two consumption peaks: one big peak in summer due to high cooling consumption and one small peak in winter as a result of heating consumption.

Industrial consumption of natural gas is defined as “natural gas used for heat, power, or chemical feedstock by manufacturing establishments or those engaged in mining or other mineral extraction as well as consumers in agriculture, forestry, and fisheries”. Also included in industrial consumption are generators that produce electricity and/or useful thermal output primarily to support the above-mentioned industrial activities.” Compared to residential, commercial and power sector of natural gas use, industrial use of gas is relatively stable, with small peaks in the winter. This sector composed 30.6% of the total natural gas consumption in 2016.

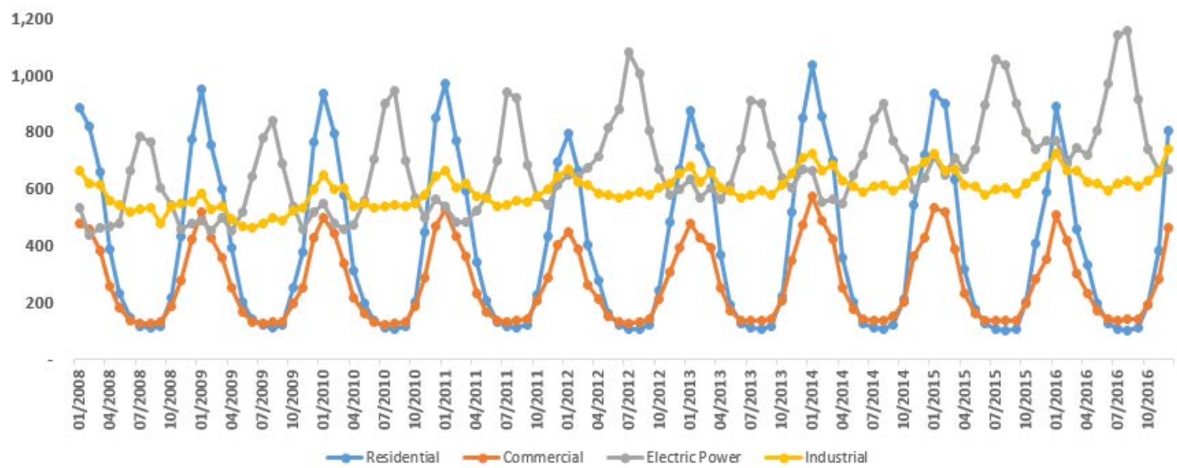


Figure 1. Monthly Natural Gas Consumption by sector (Bcf)

Figure 1 above depicts natural gas consumption in the different sectors from 2008 to 2016. Total natural gas consumption follows a strong seasonal pattern, in terms of both average and range perspective. There are two consumption peaks: a big consumption peak in winter due to the high heating consumption in residential and commercial sectors, and a small consumption peak in summer resulting from high cooling consumption, which is mainly reflected in the power sector. Averaging through 2001 to 2016, the average monthly total natural gas consumption is displayed in Figure 2. December, January, February and March have the highest consumption due to low temperature and the resulting high heating consumption. July and August have a small peak consumption because of high cooling consumption. The other months lies in between. As weather is volatile and the main driver for winter consumption, winter months' total natural gas consumption displays larger variance. In Figure 2, the vertical bar for each month indicates the consumption range from 2001 and 2016.

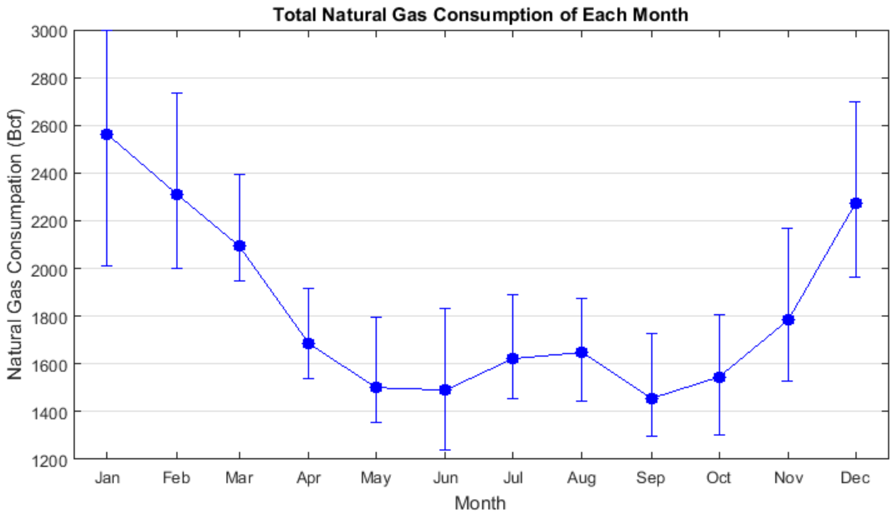


Figure 2. Average Monthly Natural Gas Total Consumption 2001-2016 (Bcf)

U.S. natural gas domestic gas production has increased dramatically since 2005. From 1973 to 2005, U.S. natural gas production was relatively stable. Since 2005, production increased significantly because of the shale gas boom facilitated by the technology advance of hydraulic fracking and horizontal drilling. In 2005, U.S. natural gas supply was 18,927 billion cubic feet with negligible gas withdrawal from shale gas wells. Supply reached 28,752 billion cubic feet in 2015 with 47% of the production coming from shale gas wells. Total production increased by 51.91% in 8 years. Unlike natural gas consumption, natural gas production does not have an obvious seasonal pattern within a year. The production is quite smooth through the year. Production in each month is just around the yearly average level. Many of the monthly fluctuations are the result of extreme weather events such as Hurricane Katrina in 2005 August and Hurricane Rita in 2009. Other than that, natural gas production throughout a year is quite stable (Figure 4).

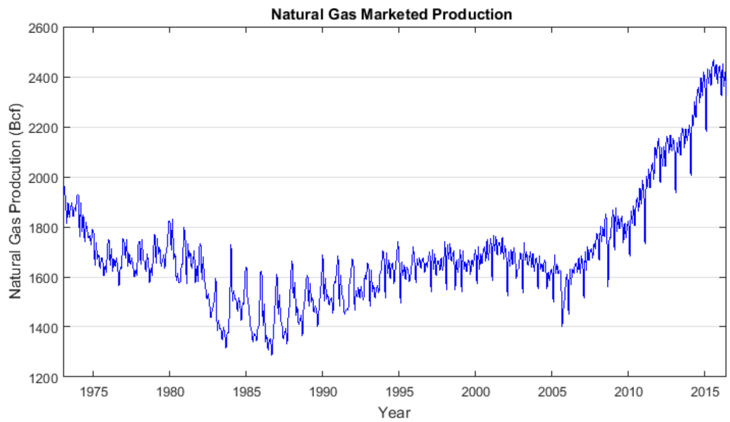


Figure 3. U.S. Natural Gas Marketed Production (Bcf)

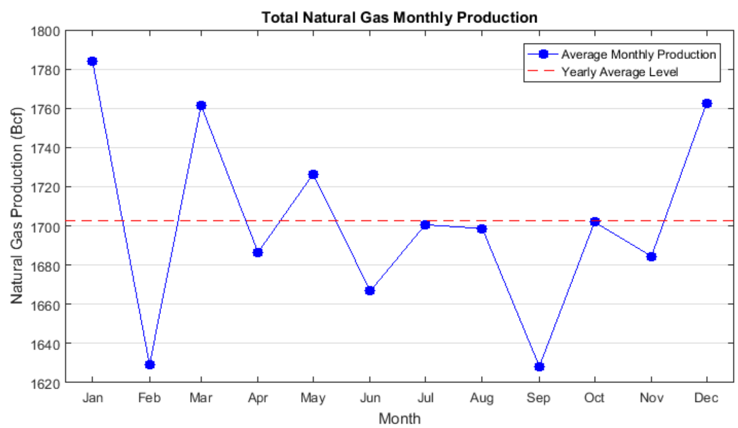


Figure 4. U.S. Average Monthly Natural Gas Production (Bcf)

The U.S. net natural gas import decreases significantly from 3.6 tcf in 2005 to 0.7 tcf in 2016. Net import decrease accounts for 31% of the total production increase of 9.4 tcf from 2005 to 2016 while domestic consumption increase uses the remaining. Until 2016, U.S. is a net natural gas importer with imports coming primarily from Canada and Mexico through pipelines, and some from Trinidad in the form of liquefied natural gas (LNG). U.S. also exports natural gas to Canada and Mexico through pipelines and to Asian countries such as Japan in the form of LNG. U.S. natural gas exports increase dramatically in the beginning of 2000's, while imports decreased since 2005, when the natural gas production increases. As a result, net import decreased substantially since 2005, as illustrated in Figure 5.

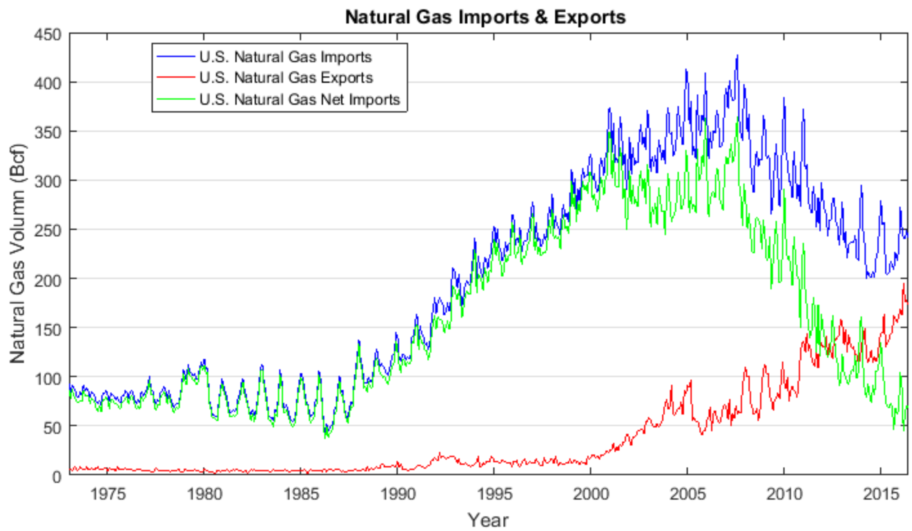


Figure 5. U.S. Natural Gas Imports, Exports from 1973 to 2016

Inventory plays an important role in natural gas market due to seasonal imbalance of production and consumption. Natural gas consumption has strong seasonal pattern while production is relatively stable all year round due to facility capacity constraints. To balance the consumption and supply of natural gas, storage is necessary and follows obvious seasonal pattern. Different months have different storage behaviors. April to October can be seen as the injection season, with very low withdrawal amounts and very large injection amounts and net storage is positive. The produced natural gas will be partly saved for winter peak season use. September and October marks the end of injection season and the total working gas storage reaches its highest level. November to March is withdrawal season, with large amounts of stored gas being used and nearly no injection. In winter, storage level is one of the major price drivers as it serves as a large portion  of total supply. For example, 971 Bcf natural gas storage was used in 2014 January, accounting for over 30% the total consumption during this period. Storage is volatile and significantly affected by consumption shifters, mainly weather factors. The huge 2014 January consumption of natural gas storage is mostly a result of extreme cold winter.

## 1.2 Natural Gas Price Behavior

Natural gas price displays an obvious downward trend in recent years especially since 2008 due to technological advancement of shale gas. As Figure 6 exhibits, from 2001 to 2008, Henry Hub natural gas spot price was generally above $5 while it remained under $5 thereafter for most of the time. This is consistent with production increases beginning in 2005 while total consumption is relatively stable. Seasonal natural gas consumption in residential, commercial and electrical sector results in seasonal natural gas price pattern: price is slightly higher in heating season. A cold winter would induce large heating consumption in residential and commercial sectors, and large consumption in electric power sector. Increased consumption puts upward pressure on natural gas price. However, the price difference between months is not exceptionally large because storage can be used to balance total supply and desired consumption.

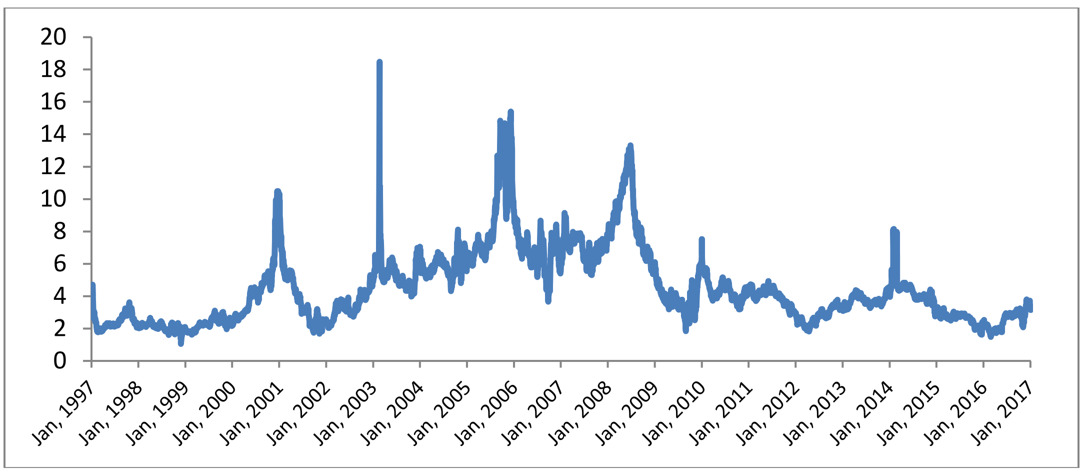


Figure 6. Henry Hub Natural Gas Spot Price ($/MMBtu)

Natural gas price is one of the most volatile among energy prices. Natural gas demand and supply in the short-term is quite price inelastic. On the consumption side, most natural gas appliances and facilities are durable goods and require considerable investment and time to change. As the major usage, heating and electricity feedstock are necessity goods so that natural gas consumption is quite rigid in the short-term. Similarly, increasing natural gas production requires long-term investment. Storage is constrained by facility capacity as well. As a result, short-term natural gas supply is inelastic. Besides the supply constraint, the transportation capacity to deliver natural gas to end users is limited by pipeline capacity and is difficult to increase in a short period. The inflexibility of both supply and consumption sides lead to abrupt price change when there is a shock to either the supply side or the demand side. This feature is especially magnified in winter heating months. As shown in Figure 7, winter season (November to March) price shows higher average level and the higher price level is consistent with higher consumption in heating season. Price is lowest for seasons with not much heating and cooling consumption. Natural gas price is more dispersed when average consumption is high, as the main driving factor – weather is quite volatile, resulting in volatile consumption. The observed historical price feature is consistent with consumption and supply features discussed above.

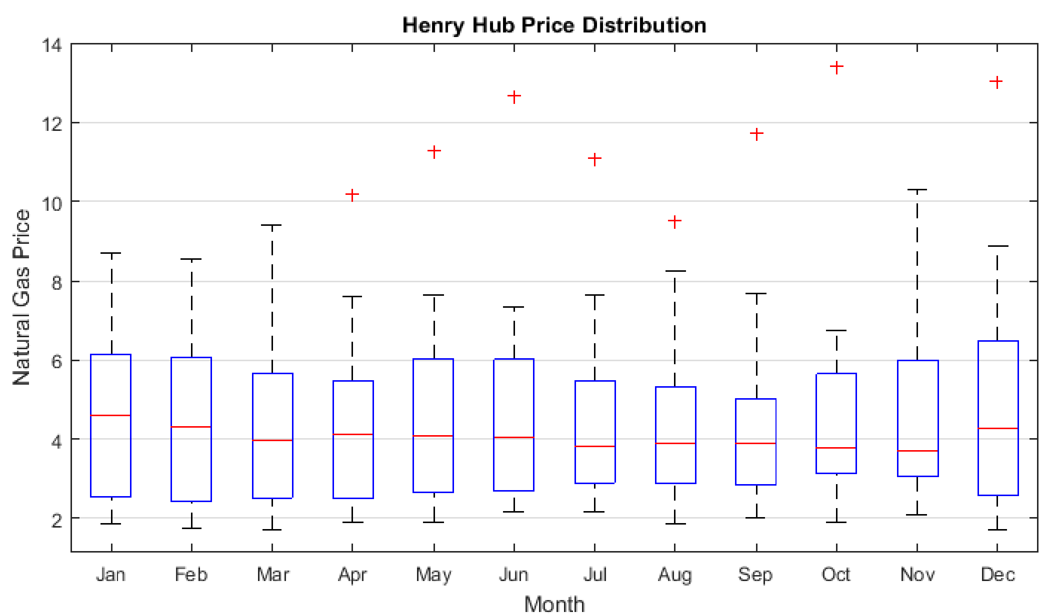


Figure 7. Henry Hub Natural Gas Price Distribution by Month – Boxplot

To better measure natural gas price volatility, Henry Hub daily natural gas spot price was collected and summarized. Price change is defined as natural log of price ratio between current and prior period: $\vartriangle p_t = \ln(p_t / p_{t-1})$. As Henry Hub daily spot price is only available on trading days, all *t* refers to trading days (excludes weekend and holiday). As summarized in Table 1, historical data shows larger volatility in winter heating months than in other months. From 1997 to early 2017, average daily price change is generally larger than 3% during October to February, when heating need is high and thus natural gas consumption peaks. The other months have smaller average price change. The standard deviation of daily price change exhibits consistent result: October to March shows much larger standard deviation of price change compared to other months. The result using data beginning in 2010 is similar: winter months' daily price change shows large average level and standard deviation. However, the magnitude is in general smaller than the whole 1997~2017 period. Both price and its volatility show seasonal patterns, which mainly relates to consumption change due to weather condition.

Table 1. Henry Hub Daily Spot Price % Change Level and Standard Deviation

| | 1997 - 2017 | | 2010 - 2017 | |
|---|---|---|---|---|
| Month | Average of Absolute Daily Price Change (%) | StdDev. of Daily Price Change | Average of Absolute Daily Price Change (%) | StdDev. of Daily Price Change |
| January | 3.40 | 4.63 | 3.26 | 4.33 |
| February | 3.53 | 6.98 | 3.53 | 6.33 |
| March | 2.51 | 4.15 | 2.76 | 5.25 |
| April | 2.04 | 2.70 | 1.92 | 2.75 |
| May | 2.14 | 2.93 | 1.96 | 3.06 |
| June | 2.17 | 3.07 | 1.80 | 2.58 |
| July | 2.15 | 2.81 | 1.52 | 2.12 |
| August | 2.34 | 3.33 | 1.43 | 1.90 |
| September | 2.95 | 4.21 | 1.67 | 2.19 |
| October | 3.47 | 5.03 | 1.79 | 2.52 |
| November | 4.11 | 6.35 | 2.89 | 3.93 |
| December | 3.40 | 5.13 | 2.57 | 4.21 |
| Grand Total | 2.84 | 4.44 | 2.25 | 3.65 |

As discussed above, weather is a key factor driving natural gas consumption, storage, price and price volatility. For example, early 2016 saw mild weather and consumption was weak compared to average level. As a result, natural gas price reached a 17-year low of $1.64. Contrast to the case of 2016, the 2013 winter was unusually cold, underground working gas storage decreased to an almost 10-year history low of 857 bcf and price hit 5-year high of $6 per MMBtu in April, 2014. Price increased sharply from October 2013 until February 2014, then reverted back to normal range gradually afterwards.

Volatile natural gas price creates difficulty for energy price risk management for all market participants. The uncertainty drives the need for hedging and volatility is a key input to value derivatives like futures contracts. Natural gas volatility spills over to other energy sectors as well. A considerable amount of natural gas is used for marginal electricity plant feedstock instead of baseload fuel. Thus, large natural gas price movement indirectly affects electricity market. Understanding natural gas price determinants, volatility and dynamics is crucial for hedging the risk associated with it.

## 1.3 Literature Review

There are many studies in the literature covering various aspects of the natural gas market: including determinants of price, consumption trend and estimation, its relationship with other energy source, and the impact of shale gas boom. With respect to the economics of natural gas storage and price volatilities, which is the focus of this research, the current natural gas storage literature focuses primarily on the valuation of a specific storage facility using different methods. A representative paper by Carmona and Ludkovski (2010) develops an "optimal switching model", using dynamic programming to choose a series of time and actions (injection or withdrawals) to maximize profit. Schoppe (2010) uses the knowledge gradient approach with non-parametric estimation method. Boogert and Jong (2008) extend the Least Square Monte Carlo method for American options to value a storage facility. Thompson (2009) solves a system of partial differential equations using the method of radial basis function collocation. Hoffler

(2007) estimates consumption for natural gas storage based on historical data extrapolation. Swing consumption (winter consumption minus summer consumption) is satisfied by indigenous production, import and storage. Optimal storage equals swing consumption minus production and import. De Joode's (2010) method is similar to Hoffler (2007), and bases the estimation on the comprehensive GASTALE model. These papers related to natural gas storage do not take weather shocks into account even though weather is the most important affecting factor of natural gas storage.

The relationship among natural gas price, weather shock and storage levels are discussed in other papers, most of which use time-series models. Geman and Ohana (2009) conduct rank correlation tests and found that natural gas inventory level and price volatility are negatively related. This negative correlation is pronounced only during periods of scarcity. The correlation increases significantly in winter for natural gas. The GARCH model is often used in understanding the relationship among natural gas price, price volatility and storage level. GARCH is able to capture the heteroskedasticity in the volatility of prices. Mu (2007) uses a GARCH model to estimate how weather shocks affect the conditional mean and variance of natural gas futures returns and finds that weather is an important factor in determining natural gas futures returns. Pindyck (2004) utilizes GARCH to analyze natural gas price volatility and finds that natural gas volatility fluctuates over time and the changes are transitory.

While the above papers do well in explaining volatility observed in financial data, there is no solid theoretical foundation that justifies their usage in natural gas markets. Economic theory argues that price changes are rooted in changes in supply and consumption factors. However, it is difficult to link this fundamental supply-consumption relationship to GARCH models. Kanamura (2009) constructed a supply and demand based volatility (SDV) model and shows that there exist inverse leverage effect in energy market, that is volatility increases in energy price. van Goor and Scholtens (2014) generalize the SDV model to account for different consumption and supply assumptions. However, this model does not take natural gas inventory into account. Weather's impact on pricing dynamics is not discussed in detail as well.

The existing literature supports the statement that market fundamentals are the key determinant of natural gas price and variance. However, most of the models used to explain natural gas prices focus on historical price realizations and do not provide insights as how to how price patterns are determined by market participants' behavior. In addition, their reliance on historical observations prohibits the models from explaining if there is structural change or shock on the consumption or supply side, and how the natural gas price and its volatility evolve over time. There is also a  large structural models named NEMS which is maintained by EIA that can analyze natural gas storage and price. However, NEMS does not  focus on the weather effect on the storage and prices, so it cannot explain the relationship between price changes and weather explicitly. More importantly, natural gas price is quite volatile compared to other energy prices and mostly driven by weather conditions. However, how price volatility is related to weather conditions has not been carefully discussed in the current literature.

## 1.4 Purpose

In general, the existing literature finds that inventory plays an important role in the natural gas market. Natural gas price and volatility display seasonal patterns that are driven by weather conditions and storage level. However, the interrelationship is not captured by a single comprehensive model and has not yet been discussed in detail in the current literature, using a model that incorporates major market participants' behavior.

This research aims to construct a model that can mimic the major market participants' behavior and reproduce the natural gas price with mean and standard deviation patterns consistent with historical observation: higher average level and standard deviation in peak consumption season. We estimate natural gas consumption and supply function in a way that the current period volume is highly correlated with previous period volume. Price will be inter-temporally correlated under competitive storage model. The objective of this study is to explain natural gas price and storage behavior based on market fundamentals, including consumption and supply.

We can utilize the framework to understand the impact of storage level and weather conditions on price levels and volatility over time. This research will quantify weather variables' impact on natural gas price and its standard deviation. In addition, we can discuss the relationship between storage and price behavior.

The thesis is organized as following: a conceptual natural gas competitive storage model is constructed in chapter two. In order to solve the conceptual model, the major functions involved like consumption, production and net storage cost functions are specified and calibrated in chapter three. Chapter four describes the numerical solving algorithm in detail. Model performance and results are discussed in chapter five. The usefulness of the model is illustrated in Chapter 6 where the impacts of increased future demand for LNG are simulated.

# CHAPTER 2. CONCEPTUAL MODEL FRAMEWORK

This chapter conducts a brief literature review of the competitive storage model and explains why using a monthly competitive storage model is suitable for analyzing natural gas price dynamics. The conceptual model framework is also presented in this chapter.

The competitive storage model is widely used for agricultural commodities and does a good job in explaining price dynamics in the short run. Williams and Wright (1991) combine a market equilibrium condition with a non-arbitrage condition to construct a rational-expectations competitive storage model for grain. Total demand and total supply clears the market while price satisfies inter-temporal non-arbitrage conditions at the same time. Miranda (1997) proposes and compares different methods to solve the competitive storage model and concludes that the collocation method performs better than alternatives. Deaton and Laroque (1992, 1995 and 1996) test the power of the competitive storage model by empirically estimating its relevance for eleven commodities. They found that the high level of serial correlation of commodity price could not be replicated by arbitrage storage behavior. To address the inability of the rational expectations storage model to explain high price correlation observed by Deaton and Laroque, Miranda and Rui (1999) assume a classical supply of storage function capturing the negative intertemporal price spreads when stocks are positive. Marginal storage cost is assumed to comprise of both physical storage cost and a marginal “convenience yield”. The application of classical supply of storage function addresses the issue examined by Deaton and Laroque. Cafiero et al. (2011) uses a finer grid when estimating the rational expectation storage model and obtain a high level of autocorrelation for major commodities as well. Peterson and Tomek (2005) apply a monthly competitive storage model to simulate U.S. corn market cash price and futures price with key characteristics consistent with historical observations.

Most applications of the competitive storage model analyze policy scenario impacts for a specific commodity or several commodities (Miranda and Glauber, 1993; Lence and Hayes, 2002). Tran (2013) utilizes the model to assess the food price impact resulted from increased corn consumption from US ethanol production and mainly focused on its price impact in transitionary period.

The rational expectation storage model has not been applied to the natural gas market to my knowledge. One question that needs to be answered is whether the model is actually suitable for explaining natural gas market and price dynamics. As discussed above, storage plays an important role in balancing seasonal consumption and production. The existing literature illustrates the significant causal relationship between inventory and natural gas price. With storage's importance established, the model would apply if natural gas price can be mostly explained by market fundamentals like consumption, production and storage behavior instead of simple speculation. Hulshof et al. (2016) concludes that natural gas price is mainly affected by gas market fundamentals including availability and temperature. Nick and Thoenes (2014) using a structural VAR model shows similar results: natural gas price is mainly driven by temperature and total supply in the short term. Knittel and Pindyck (2016) use the storage model to analyze the crude oil price spike around 2008 and conclude that the price change is consistent with production, inventory and convenience yields. Speculation had minimal effect on crude oil prices. Using structural vector autoregression (SVAR) model, Wiggins and Etienne (2017) reach a similar conclusion for natural gas market as Knittel and Pindyck (2016) did for crude oil markets. Their SVAR model indicates that during 1993 to 2015, natural gas price fluctuation is mainly driven by supply and consumption shocks. Speculative activities plays a minor role. Based on the existing literature, we believe that rational expectations storage model which incorporates the behavior of major market participants' behavior is theoretically applicable for explaining natural gas market.

As natural gas is produced continuously, there are natural gas injection and withdrawal every month. Natural gas consumption for residential, commercial and electrical sectors have obvious seasonal pattern. We thus employ a monthly model instead of an annual or quarterly one.

This study follows Williams and Wright's (1991) concept of the competitive storage model: price and storage are determined so that natural gas total supply and total consumption clears the market each month. Natural gas price satisfies an inter-temporal non-arbitrage condition at the same time. The competitive storage model presented here simplifies the natural gas market while incorporating the main behavior of key market players. A single competitive national natural gas market is assumed to exist. Three major agents in the natural gas market are modeled: natural gas producer, gas consumers and storer. Producers drill gas wells and extract natural gas to deliver to the market. Four broad categories of gas end-users including residential, commercial, industrial and electrical power consumers consume gas. The storer holds natural gas to future periods as long as the price appreciation covers the net cost of carry. For each period, the market clears so that the total supply of natural gas equals total gas consumed and stored. Monthly model is used to better capture the seasonal variation of key seasonal measurements like price, consumption and storage.

## 2.1 Consumption

As discussed above, natural gas end-use consumption has four broad categories: residential, commercial, industrial and electric power sector. So the four most dominant sectors of natural gas consumption considered in the storage model are natural gas consumption of residential ($D_r$), commercial sector ($D_c$), electric power sector ($D_e$) and industrial sector ($D_{in}$). The total end-use natural gas consumption is the summation of the consumption of these four sectors.

$$D_{tot}^{(m)} = D_r^{(m)} + D_c^{(m)} + D_e^{(m)} + D_{in}^{(m)}, \quad (1)$$

where *m* denotes the month, which corresponds to January, February, …, December.

Natural gas consumption in each month is mainly driven by natural gas price and weather. More specifically, besides the impact of price on consumption, the consumption of all four sectors generally increases with cold weather conditions for heating purpose. In summer, the electric sector is highly driven by cooling demand.

To evaluate weather impact on natural gas consumption, two variables are used to reflect energy consumption under a specific temperature: heating degree days (HDD) and cooling degree days (CDD). By the convention of National Oceanic and Atmospheric Administration (NOAA)[1],

$$HDD = max(0{,}65 - T), \quad CDD = max(0, T - 65),$$

where T denotes the daily temperature in degrees Fahrenheit.

[1] National Weather Service, The National Oceanic and Atmospheric Administration (NOAA).

http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/ddayexp.shtml

The state level HDD and CDD are directly available from NOAA. For regional level HDD and CDD, we use the monthly gas-population weighted cooling degree days and heating degree days. The monthly degree days are the summation of the degree days within that month.

Monthly natural gas consumption are highly correlated with cooling degree days and heating degree days (*cdd* and *hdd*). With the consideration of price impact, natural gas consumption for different sectors $D_i$ is generally a function of price and weather condition.

$$D_i^{(m)} = D_i^{(m)}\left(p^{(m)}, hdd^{(m)}, cdd^{(m)}\right),\ D_i \in \{D_r, D_c, D_e, D_{in}\}.$$

The total monthly consumption is the summation of consumption from these four sectors, and is accordingly a function of price (*p*), heating degree days (*hdd*), and cooling degree days (*cdd*). As consumption is subject to a random shock ($\varepsilon_D^m$), the total consumption of each month can be expressed as:

$$D_{tot}^{(m)} = D_{tot}^{(m)}(p^{(m)}, hdd^{(m)}, cdd^{(m)})\varepsilon_D^m. \tag{2}$$

## 2.2 Production

The producer's decision is made based on three factors: production cost, expected next month price and expected next year price. Production is determined by short-term planning as well as medium to long-term investment of drilling and extraction facilities. It is assumed that natural gas production using existing wells takes one month from planning to realization. The expected next year price represents investment incentives for producers. Producers are assumed to react to market price change one year ahead. This specification is different from consumers, whose consumption is based on current observed gas price.

To be more specific, the production decision is made in month m. In month m, period (m+1) price is unknown and producers can only respond to expected next month (m+1) price $E_m\left(p^{(t,m+1)}\right)$. Producers are sophisticated market agents. Their investment activities (e.g.

drilling extra wells) are forward-looking, based on cost and next year price expectation to respond to price change due to market change. The investment starting in year (t-1) is dependent on average price in year t and will be effective in year t. The realized gas production is subject to random shock due to uncertainties like productivity and production suspension resulting from extreme weather conditions. Natural gas production is determined as following:

$$Pr^{(t,m)} = \beta_0 Pr^{(t,m-1)} \left( \frac{E_{t-1}(P_{mean}^{(t)})}{cost^{(t)}} \right)^{\beta_1} \left( \frac{E_{t,m}(p^{(t,m+1)})}{E_{t-1}(P_{mean}^{(t)})} \right)^{\beta_2} \varepsilon_{pr}, \qquad (3)$$

where t denotes year and m represents month; $E_{t-1}(P_{mean}^{(t)})$ is the expected consumption-weighted average price of all 12 months of year (*t*) at the beginning of year (t-1); $E_{t,m}\left(p^{(t,m+1)}\right)$ denotes the expected price next month; $cost^{(t)}$ is the production cost of year t; $\beta_1$ and $\beta_2$ represents the investment and short-term price elasticity respectively; and $\varepsilon_{pr}$ is the production shock.

When the mean price of a year is higher than the cost, the producers tend to increase the investment and produce more natural gas, and vice versa.

The monthly total supply ($TS^{(m)}$) of natural gas is composed of the previous month total production of the natural gas, net imports, and the current total working gas storage.

$$TS^{(m)} = Pr^{(m-1)} + netIm^{(m-1)} + s^{(m-1)} = TP^{(m-1)} + s^{(m-1)}, \qquad (4)$$

where $TP^{(m-1)}$ denotes the total supply from the producer including the monthly production of previous month ($Pr^{(m-1)}$) and the monthly net import ($netIm^{(m)}$). $s^{(m-1)}$ denotes the storage carried over from the previous month. As net import volume is small relative to total production and consumption level, it is ignored in this section.

According to equation (4), the total production in the current month serves as the natural gas supply of next month. Taking into account random shock production, the total production of the month thus can be expressed as:

$$TP^{(m)} = TP\Big(E_{t,m}(p^{(t,m+1)}), E_{t-1}(P_{mean}^{(t)})\Big)\varepsilon_{Pr}^{m}. \tag{5}$$

2.3 Equilibrium Condition

The natural gas available at month $m$ will be either consumed immediately or stored for future use. The market clears so that monthly total supply equals monthly total consumption and the storage for next period. The storage can be seen as speculative or precautionary consumption. That is: $TS^{(m)} = D_{tot}^{(m)} + s^{(m)}$. With equations given above, we have

$$TS^{(m)} = TP^{(m-1)} + s^{(m-1)} = D_{tot}^{(m)}\big(p^{(m)}, hdd^{(m)}, cdd^{(m)}\big) + s^{(m)}, \tag{6}$$

Or equivalently,

$$s^{(m)} = TS^{(m)} - D_{tot}^{(m)}\big(p^{(m)}, hdd^{(m)}, cdd^{(m)}\big)$$

$$= TP^{(m-1)} + s^{(m-1)} - D_{tot}^{(m)}\big(p^{(m)}, hdd^{(m)}, cdd^{(m)}\big) \geq 0,$$

Storage is the remainder of the total availability minus the amount of natural gas consumed.

2.4 Storage Decision

Natural gas storage facilities are now operated in a very competitive way. Prior to 1994, pipeline companies own most of the natural gas storage and had exclusive control of the storage facility utilization. Federal Energy Regulatory Commission (FERC) Order 636 requires interstate pipeline companies to operate storage facilities on an open-access basis. That is, “the major portion of working gas capacity (beyond what may be reserved by the pipeline/operator to maintain system integrity and for load balancing) at each site must be made available for lease to third parties on a nondiscriminatory basis.” After natural gas market deregulation and liberalization, natural gas storage facilities can be rented and used by any participants with the sole purpose of profit maximization. (De Jong and Walet, 2003)

As the natural gas market now can be seen as competitive market, storage level is determined by arbitrageurs. We assume the arbitrageurs are rational decision makers. From period t to period t+1, inventory holders want to maximize their profit while running the storage facility. Assuming a constant monthly interest rate *r* and denoting the total net storage cost as *sc*, the storer's maximization problem is as follows.

$$\max_{s^{(m)}} E\left(\pi^{(m+1)}\right) = \frac{1}{1+r} E_m(p^{(m+1)}) s^{(m)} - p^{(m)} s^{(m)} - sc^{(m)}, \qquad m = 1, 2, 3, \ldots, 12, \quad s.t.\ s^{(m)} \geq 0, \tag{7}$$

where $E_m\left(p^{(m+1)}\right)$ is the expected price of natural gas of the next month (m+1, when m = 12 (Dec), m+1 = 1 (Jan)). $sc^{(m)}$ is the total net carry cost of the current month. $r$ is the discount rate, we use LIBOR 1-month interest rate in this study.

The first order condition is

$$\begin{cases} \frac{1}{1+r} E\left(p^{(m+1)}\right) - p^{(m)} - {sc'}^{(m)} = 0, s^{(m)} > 0, \\ \frac{1}{1+r} E\left(p^{(m+1)}\right) - p^{(m)} - {sc'}^{(m)} < 0, s^{(m)} = 0, \end{cases}$$

where $sc'$ is the marginal net carry cost.

Alternatively, we can rewrite above conditions in the format to define storage demand:

$$\begin{cases} \left(\frac{1}{1+r} E\left(p^{(m+1)}\right) - p^{(m)} - {sc'}^{(m)}\right) s^{(m)} = 0, \\ \frac{1}{1+r} E\left(p^{(m+1)}\right) - p^{(m)} - {sc'}^{(m)} \leq 0. \end{cases} \tag{8}$$

Intuitively, arbitrage inventory holders store natural gas in the anticipation of profit. If the net storage cost of one unit of natural gas is higher than the expected profit from holding stock from period m to next period (m+1), no gas will be stored. When cost of storage is low and there is profit of inventory, storers will buy more gas from the market with purpose of selling it at a higher price in the future.

Natural gas consumption is seasonal and most of the storage will be carried over to winter to satisfy peak natural gas usage from other seasons. The above equation defines the contemporary non-arbitrage condition on a monthly basis. That is, the non-arbitrage is based on one month ahead expectation. To validate that the above formulation is suitable to model the natural gas storage behavior, we need to justify that a one month ahead arbitrage equation is sufficient to capture seasonal carryover of gas storage.

In an ideal market, the contemporary non-arbitrage condition holds for any kind of time frequency defined, whether it is a long time horizon such as one year or more frequent intervals like weekly or daily, if the market is liquid enough and there is no transaction cost or friction. Otherwise, speculative inventory holders will seek profit from trading until the non-arbitrage condition is met.

The non-arbitrage condition is applicable not only for speculative storage but also applies to precautionary storage. For speculative storage, the sole purpose is to buy at a low price, hold inventory and sell it at higher price in next period. For precautionary inventory holder, mostly power plant and manufacturers, they store natural gas for convenience purpose. The storage can reduce their cost to revise their production schedule and decrease the risk of stock out. To incorporate both speculative and precautionary storage into the model, we define the net storage cost as a combination of physical storage cost and convenience yield. The natural cycle of natural gas storage is one year, with gas injection in late spring to early fall when natural gas production is higher than the volume consumed, and gas withdrawal in winter to meet high heating need. The yearly cycle is incorporated into the model with the notion of “normal storage level”. Normal storage level will be used for net storage cost calibration. More details will be given in chapter three “model specification and calibration” section.

With the development and expansion of high-deliverability storage sites, natural gas storage facilities are not restricted to seasonal cycling and are able to serve the purpose of short-term profit maximization. Most natural gas is stored underground in three major types: depleted fields, salt caverns and aquifers. Among the three types, salt caverns have the highest deliverability and highest cycling rate. A salt cavern has the ability to perform as many as a dozen withdrawal and injection cycles within one year. They are very responsive to consumer consumption and price change as salt cavern storage field is able to send out natural gas in a very short time, even within an hour. In addition, gas withdrawal and injection can happen at the same time for salt cavern field. As of 2015, the storage capacity of salt cavern reaches 709Bcf, increasing from 189Bcf from 2000. The increasing number and high capacity of salt cavern makes high frequency arbitrage feasible.

Monthly instead of a yearly or quarterly model is used in this study as consumption and its key drivers varies even within same season. The employment of a monthly model will better capture the monthly consumption variation. Monthly model is able to quantify weather impact more accurately than seasonal models. In addition, the purpose of this study is to generate monthly price series that is consistent with observed prices.

To summarize the timeline of decision making of the three major market participants discussed above: Production decision is made in period (m-1) and the actual produced amount is realized in period m with shock $\varepsilon_{Pr}^{m}$. At the end of period (m-1) or equivalently at the beginning of period m, the stock carried over from last period is a known number $s^{(m-1)}$. In period m, weather variables HDD, CDD and consumption shock $\varepsilon_D^m$ will be realized. State variables denote the available information of each month for decision making. The state variables in each months are: total supply $TS^{(m)}$ with supply shock $\varepsilon_{Pr}^m$, consumption shock $\varepsilon_D^m$, and weather conditions HDD and CDD.

The purpose of the conceptual model is to solve for a policy function $p^{(m)}$ so that 1) the market clears at price $p^{(m)}$ given the format of total consumption function ($D_{tot}^{(m)}$), and 2) inter-temporal non-arbitrage condition is satisfied through the control of amount of natural gas stored for next period usage.

Chambers and Bailey (1996) proved the existence of a unique rational expectations equilibrium. We define price as a function of state variables: weather condition, total supply and consumption shock.

$$p^{(m)} = p^{(m)}\left(hdd^{(m)}, cdd^{(m)}, TS^{(m)}, \varepsilon_D^m\right). \tag{9}$$

To account for weather shock, we see HDD and CDD in each month as random variables with certain distributions. Thus, we can substitute the above equations to the first order condition and get the below equation when stock is positive.

$$\frac{1}{1+r} E\left(p^{(m+1)}(hdd^{(m+1)}, cdd^{(m+1)}, TP^{(m)} + s^{(m)}, \varepsilon_D^{m+1})\right) \tag{10}$$
$$= p^{(m)}\left(hdd^{(m)}, cdd^{(m)}, TS^{(m)}, \varepsilon_D^m\right) + sc'^{(m)}\left(s^{(m)}\right).$$

When we substitute the price function and take the price expectation with respect to the weather conditions ($hdd^{(m+1)}$ and $cdd^{(m+1)}$) and random consumption and supply shocks, we would observe the storage decision as an endogenous function of price, showing the relationship of the monthly storage level ($s^{(m)}$) to the current weather conditions ($hdd^{(m)}$ and $cdd^{(m)}$) and the total supply ($TS^{(m)}$). Subsequently, we can make storage decision based on the current state variables as well (weather conditions and total supply).

$$s^{(m)} = s\left(hdd^{(m)}, cdd^{(m)}, TS^{(m)}, \varepsilon_D^m\right). \quad (11)$$

Equation (2), (3), (6) and (8) together defines the competitive storage model framework.

# CHAPTER 3. MODEL SPECIFICATION AND ESTIMATION

To solve the rational-expectations storage model and apply the model to scenario analysis, the functions need to be calibrated to mimic the behavior of producers and consumers as close as possible. Regarding the scope of this study, we need to calibrate 1) the natural gas demand functions to capture consumer’s response to price and weather condition changes; 2) the natural gas supply function to fit and predict producers’ decision based on their expectations about natural gas price; 3) storage cost and convenience yield because storage facility operators make decisions based on current price, expected price and the cost to carry natural gas from current period to the next period. Storage cost and convenience yield is an integral part of their decision making process.

This chapter describes the function specification, calibration methodology and discusses the estimation result.

## 3.1 Demand Specification and Estimation

Identifying natural gas demand and supply functions is important not only in the sense of them being key inputs into the competitive storage model of this study, but also indicative of understanding of market participant’s behavior and their differences among sectors. There are many existing studies estimating natural gas consumption and supply elasticities. However, the number of studies using recent data to capture the structural change due to significant production increases since 2005 is limited. In addition, most studies focus on one consumption sector or supply only. A full set of consumption estimates covering all sectors, including residential, commercial, industrial and electrical using the most recent data and consistent methodology is in need. In this way, the elasticities for different sectors will be comparable and capture sector differences, in terms of the magnitude of responsiveness to price change and dependence on weather conditions.

All natural gas demand functions are assumed to follow AR(1) processes. As illustrated by Deaton and Laroque (1995, 1996), the serial correlation in commodity prices can be largely explained by supply and consumption autocorrelation in addition to speculative storage behavior. Consumption is sticky and any adjustment to consumption habit is made gradually. When there is a structural change in the supply-consumption relation, such as LNG export, or technology evolution that introduces higher consumption or supply, the inclusion of prior value will help ensure gradual market change, instead of a sudden change responding only to price.

For all sectors, consumption is calibrated by season instead of applying a yearly consumption function to better capture the impact from weather variables and responsiveness to price movement in different seasons. Taking the residential sector for example, natural gas is mainly used for space heating in winter. The consumption level is driven primarily by weather conditions and HDD is highly positively correlated with consumption level. While in summer, most residential consumption comes from cooking, which is relatively insensitive to temperature change. Consumers do respond to different extents to the same driving factors throughout the year. As one of the major purpose of this study is to assess how much weather conditions can explain price change, it is important to accurately quantify the relationship between consumption and weather variables during the calibration process.

Consumption is calibrated at the regional level. This study utilizes the U.S. climate region breakdown as defined by NOAA.[2] The continuous United States is divided into 9 climate

---

[2] Source: NOAA website https://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php#references

regions: Northeast, Southeast, Ohio Valley (Central), Upper Midwest, South, Northern Rockies and Plaines, Northwest, Southwest and West. Hawaii and Alaska are excluded. The area within each region is climatically consistent (Karl and Koss, 1984).

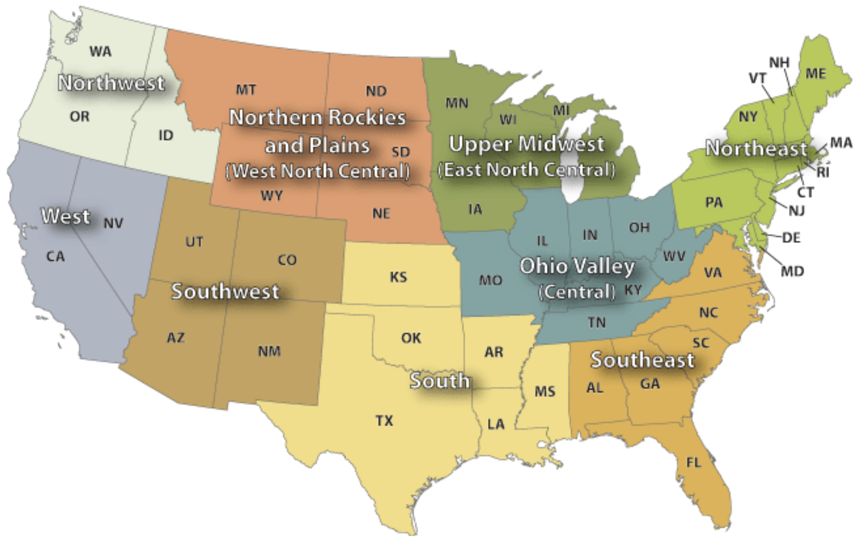

Figure 8. U.S. Climate Regions

Constant elasticity function format is assumed. To be consistent with historical observations, time trend and region dummies are added into the consumption functions as necessary. The consumption of each sector is different between months and region. Monthly dummies are added to capture monthly variations which cannot be explained by weather conditions and price movements. For example, specific calendar-driven factors, like holidays or special season, are important and will lead to month-to-month difference in consumption (Sánchez-Úbeda and Berzosa, 2007). Region specific dummy variables are added because there are significant regional difference in natural gas consumption. For northern states like Michigan and Minnesota, natural gas consumption is quite high especially in winter due to necessary heating consumption. Natural gas availability for different geographical regions varies and infrastructure is not the same as well. In addition, region-level policy affects the energy-efficient technology utilization. All these factors lead to region-to-region differences in natural gas consumption. To account for these effects, monthly dummy and region dummy are introduced to the model to capture the monthly and regional variation in natural gas consumption.

As discussed above, the regional natural gas consumption function for sector $i$ and region $j$, is expressed further as below and follows AR(1) process.

$$ln\left(D_{ij}^{(m)}\right) = \begin{array}{c} \alpha_{ij,1} * ln\left(p^{(m)}\right) + \alpha_{ij,2} * hdd_j^{(m)} + \alpha_{ij,3} * cdd_j^{(m)} \\ + \alpha_{ij,4} * ln\left(D_{ij}^{(m-1)}\right) + \alpha_{ij,5} * yr + \alpha_{i,6}^{(m)} + \alpha_{ij,7}, \end{array} \quad i = 1 - \ 4; j = 1 - 9,$$

where $D_{ij}^{(m)}$ is the monthly consumption for sector i and region j, which is negatively correlated with $p^{(m)}$ and positively correlated with $D_{ij}^{(m-1)}$, $(\alpha_{ij,1} < 0, \alpha_{ij,4} > 0)$. $p^{(m)}$ denotes natural gas price, Henry Hub natural gas spot price is used in this study; $D_{ij}^{(m-1)}$ is consumption of prior period; $hdd_j^{(m)}$ and $cdd_j^{(m)}$ are current month HDD and CDD values; time trend is expressed in year $yr$; $\alpha_{i,6}^{(m)}$ is the monthly dummy to capture monthly deviation; and $\alpha_{ij,7}$ is the region dummy variable for each consumption sector to account for region over region difference.

The total consumption can be expressed as the summation of the local consumption of all sectors and all regions, and thus as a function of price, weather conditions and prior consumption level. All the coefficients, $\alpha$'s, can be estimated from the historical observations of all consumption sectors, corresponding weather conditions and price.

$$D_{tot}^{(m)} = \sum_{i=1}^{4} \sum_{j=1}^{9} D_{ij}^{(m)} = \ D_{tot}^{(m)}\left(p^{(m)}, hdd^{(m)}, cdd^{(m)}, D^{(m-1)}, yr, \alpha^{(m)}\right),$$

here the $\alpha^{(m)} = \left\{\alpha_{ij}^{(m)}\right\}$ is the month-specific coefficients including the monthly dummy and region dummy, $D^{(m-1)} = \left\{D_{ij}^{(m)}\right\}$ is the consumption from previous month for each region, and $p$, $HDD$ and $CDD$ are all monthly explanatory variables. The coefficients of the variables are assumed different from season to season, i.e. the marginal effect of each explanatory variable is different for different seasons and constant over time.

All the elasticities are estimated using historical observations. The data source used for natural gas consumption and supply estimation is from the Energy Information Administration (EIA) and weather data is from NOAA. Both natural gas data and weather data are available on a monthly basis. The data is provided at state, regional and national level. State level data is compiled into climate region level for consumption estimation. National data is used for supply estimation. Real price instead of nominal price is used in this study and CPI data is obtained from US Bureau of Labor Statistics.

We assume price elasticity of all regions is the same. Therefore, the above consumption function can be expressed as:

$$ln\left(D_{ij}^{(m)}\right) = \alpha_{i,1}^{(m)} * ln\left(p^{(m)}\right) + \alpha_{i,2}^{(m)} * hdd_j^{(m)} + \alpha_{i,3}^{(m)} * cdd_j^{(m)} + \alpha_{i,4}^{(m)} * ln\left(D_{ij}^{(m-1)}\right) + \alpha_{i,5}^{(m)} * yr + \alpha_{i,6}^{(m)} + \alpha_{ij,7}\ , \qquad i = 1, \ldots, 4; j = 1, \ldots, 9,$$

Please note that one single national price is used for all four sectors and all regions in this study. Henry Hub price without regional or sector adjustment is used as the representative national price although natural gas price and its variation for different sectors and regions are not exactly the same. The price for residential use is most volatile and price for electric power and industrial sectors are relatively stable. The behavior of different natural gas end users varies slightly. Electrical, industrial and city gate prices adjust to market disequilibrium quickly while commercial and residential prices adjust slowly (Mohammadi, 2011). However, U.S. natural gas market is integrated and competitive after deregulation. Prices from different sectors are highly correlated. For price responsive end users like industrial and electrical sector, the correlation between the price delivered to end users and Henry Hub price is larger than 95%. The state level natural gas city gate price links closely to Henry Hub price as well, with average correlation at 91%. Henry Hub price is the most representative spot price in U.S. as the largest natural gas trading hub. Residential, commercial, industrial and electrical prices are always determined by adding a premium or discount to the Henry Hub spot price. Also, NYMEX natural gas futures price is closely related to Henry Hub price. Thus, Henry Hub price and NYMEX futures price are used in this study.

The obstacle of using historical data to estimate the models is endogeneity due to simultaneous equations bias, as only the equilibrium prices and quantities are available from historical data. There are different methods used to estimate elasticity in existing literature. Arora (2014) estimates US natural gas consumption and supply price elasticity for both short run and long run using vector-autoregression (VAR) model. The data used in the study ranges from 1993 to 2013. The estimated consumption elasticity is -0.24 and the price elasticity of supply is between 0.1 and 0.42. Another set of studies use different instrument variables for elasticity estimation with the weather variable most commonly used. Davis and Muehlegger (2010) estimates natural gas consumption by sector, including residential, commercial and industrial, using weather variables as price instrument. Hausman and Kellogg (2015) utilize similar methodology as Davis and Muehlegger, using lagged weather in other states as instrument variable.

This study follows the methodology used by Hausman and Kellogg (2015): natural gas price is instrumented by lagged weather variables in other regions. The estimation framework is similar to Roberts and Schlenker (2013): for storable commodities, storage connects prices among periods. As a result, consumption elasticity can be identified using past shocks. To avoid inconsistent coefficient estimation due to endogeneity, instrument variables are used. For consumption elasticity estimation, the instrument variable should be selected so that it can only affect consumption indirectly from price. A natural candidate for instrument variable is lagged weather in other regions: Lagged weather in other regions is correlated with natural gas price because previous months' weather drives consumption and thus storage level. From a competitive storage perspective, storage level affects price. For example, if previous months see extremely cold weather, storage level will decrease as more natural gas will be used for heating purpose. Lower storage level will increase price as total supply becomes tighter. Meanwhile, lagged weather in other regions won't directly impact natural gas consumption if the current weather of this region is controlled for.

Long run elasticity assumes energy related appliance, capital stock and infrastructure can respond to price changes. There are mainly two commonly used methodologies. The first is to incorporate capital stock utilization in the consumption estimation (Fisher and Kaysen, 1962; Dubin and Macfadden, 1984). The second is flow adjustment model or partial adjustment model (PAM). Due to limited availability of natural gas appliance and capital stock data, this study employs flow adjustment model to infer long-term price elasticity. Long run elasticity is defined based on Houthakker's (Houthakker, Verleger and Sheehan, 1974) flow adjustment methodology. This methodology is still actively used in current literature (Lin and Prince, 2013; Dahl and Sterner, 1991). Consumption in the short run is less responsive to price than in the long run. In the short run, consumers tend to stick to current habits. In addition, it takes time to invest in energy related equipment/facilities and switch into new technology. For example, it may take several years for a power plant to switch from coal to natural gas as the feedstock. In Houthakker's flow adjustment model, assuming there is no constraint, a desired consumption $q_{it}^{*}$

is assumed and it is a function of price $p_{it}$ and other variables $y_{it}$. $q^*_{it} = \alpha p^{\gamma}_{it} y^{\beta}_{it}$. The flow adjustment is set as:

$$\frac{q_{it}}{q_{it-1}} = (\frac{q^*_{it}}{q_{it-1}})^{\theta},$$

where 0< θ <1 and θ denotes the speed of adjustment. The equation can be rewritten as

$$lnq_{it} = \theta ln\alpha + \theta\gamma lnp_{it} + \theta\beta lny_{it} + (1-\theta)lnq_{it-1}.$$

The implied long-run price elasticity is $\gamma = \theta\gamma/(1-(1-\theta))$ , which can be inferred from the AR(1) coefficient and estimated short-run price elasticity.

In this study, the estimated short-term price elasticity is $\alpha_1$. The long-term elasticity will be inferred as $\alpha_1/(1-\alpha_4)$. Long-run elasticity is embeded in the consumption function specification.

Consumption price elasticity is estimated by sector, including residential, commercial, industrial and electrical sectors as each sectors' end user adjusts to price at different speed and to various extent. The regression is conducted at climate regional level to account for regional difference. As discussed above, lagged weather in other region is used as instrument variable for natural gas price. Specifically, for each region *i* at month *t*, the price instrument variable is constructed as following: get the heating degree days (HDD) for all other climate region divisions, average those HDD weighted by natural gas user population, sum the weighted average HDD from period t-1 to t-12. Instead of using previous one-month weather variable, 12 preceding months' weather is used because its impact on storage is accumulative and all months' weather deviations will be captured in this way.

The price, consumption, production, storage and weather data used in the elasticity estimation are all summarized in Table 2. January data represents winter season while July represents summer season. Consumption estimation is based on regional level so that consumption for different sectors are summarized at regional level. Weather variable: HDD and CDD listed in Table 2 are at regional level as well. Price, production and working gas in storage is at national level.

Table 2. Statistics Summary of the Major Variables used for Consumption Estimation

| | January | | July | |
|---|---|---|---|---|
| | Average | Std. Dev. | Average | Std. Dev. |
| Residential Consumption (MMcf) | 98,183 | 71,335 | 12,462 | 8,261 |
| Commercial Consumption (MMcf) | 53,817 | 40,355 | 14,142 | 9,584 |
| Industrial Consumption (MMcf) | 73,185 | 70,161 | 61,117 | 66,917 |
| Electrical Consumption (MMcf) | 63,243 | 60,005 | 101,205 | 93,035 |
| HDD | 953 | 324 | 9 | 19 |
| CDD | 1 | 2 | 305 | 136 |
| Henry Hub Spot Price ($/MMBtu) | 3.76 | 1.28 | 3.62 | 0.77 |
| NYMEX Futures Price ($/MMBtu) | 3.69 | 1.21 | 3.59 | 0.76 |
| Marketed Production (MMcf) | 1,998,108 | 281,269 | 2,046,661 | 281,363 |
| Underground Working Storage (MMcf) | 2,404,583 | 329,405 | 2,896,930 | 277,794 |

For the electrical power sector, the price ratio of natural gas to coal is used to account for the substitution between coal and natural gas. Natural gas consumption for electrical section increased significantly in the most recent decade. Coal to gas switching is primarily driven by less expensive natural gas compared to coal used for power generation. The price of coal delivered to power sector comes from EIA as well.

Two Stage Least Square estimation technique is used. The estimated result for consumption is summarized in Table 3. We mainly focus on the price elasticity and the coefficient of weather variables. For residential sector, the short-term consumption price elasticity is -0.07 in winter and consumers barely respond to price change in the summer. Residential consumers mainly use natural gas for cooking all year round and heating in winter. If price increases, consumers may reduce their usage in winter, while there is little room for usage reduction in summer. Weather is the main driver for residential gas use in winter, with a coefficient of 0.78. In the shoulder season which includes both spring and fall seasons, the coefficient of HDD is 0.41. Commercial sector behaves similarly to the residential sector. The price elasticity is -0.04 in winter and the HDD coefficient is 0.55. Industrial and electrical sector show higher elasticity than residential and commercial sector, with price elasticities at -0.06 and -0.15 respectively for winter, and -0.11 and -0.31 for summer. Weak instrument hypothesis is rejected at 1% confidence level for almost all the consumption function estimation, indicating the instruments are not weak. Table 5 to Table 8 displays the detailed estimation for each consumption sector. The first stage estimations are in Appendix table A1a to table A1d.

Table 3. Consumption Function Estimation

| | Residential | | | Commercial | | |
|---|---|---|---|---|---|---|
| | Winter | Summer | Shoulder | Winter | Summer | Shoulder |
| log(Price) | -0.07** | -0.00 | -0.03 | -0.04* | -0.04** | -0.02 |
| log(D_t-1) | 0.33*** | 0.52*** | 0.51*** | 0.33*** | 0.53*** | 0.50*** |
| Implied long-term elasticity | -0.10 | -0.00 | -0.06 | -0.06 | -0.09 | -0.04 |
| HDD | 0.78*** | | 0.41*** | 0.55*** | | 0.21*** |
| CDD | | | | | | |
| Weak Instrument | *** | *** | *** | *** | *** | *** |
| Wu-Hausman | ** | | | * | | |
| Adjusted R-squared | 0.9922 | 0.9878 | 0.9828 | 0.9928 | 0.9904 | 0.9825 |
| p-value | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 |
| | Industrial | | | Electrical | | |
| | Winter | Summer | Shoulder | Winter | Summer | Shoulder |
| log(Price) | -0.06*** | -0.11*** | -0.04 | -0.15* | -0.31*** | -0.21** |
| log(D_t-1) | 0.68*** | 0.77*** | 0.73*** | 0.80*** | 0.31*** | 0.77*** |
| Implied long-term elasticity | -0.19 | -0.49 | -0.16 | -0.75 | -0.46 | -0.92 |
| HDD | 0.10*** | | 0.01* | 0.17** | | |
| CDD | | | | | 0.64*** | 0.05** |
| Weak Instrument | *** | | *** | *** | *** | *** |
| Wu-Hausman | ** | | | | * | |
| Adjusted R-squared | 0.9965 | 0.9984 | 0.9968 | 0.9795 | 0.976 | 0.9723 |
| p-value | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 | < 2.2e-16 |

Note: 1) For residential and commercial section, the data ranges from 1997 to 2016; Industrial and electric power section data starts from 2006 to 2016;
2) Some regions have 0 bcf natural gas consumption for electrical sector in some months. For those observations, 1 bcf is used instead of 0 bcf. Result would be the same if those observations are excluded;
3) Significant codes: *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.
4) Weak instrument: Weak instrument test is an F-test in the first stage estimation. The null hypothesis is that the model has weak instrument. A rejection means the instrument is not weak.
5) Wu-Hausman test: Wu-Hausman test is used to test the consistency of the OLS estimate, assuming the IV is consistent. The null hypothesis is that the OLS estimate is consistent. A rejection of the null hypothesis means that the OLS estimate is not consistent and endogeneity exist.

Our estimation result is generally in line with the existing literature: industrial and electrical sector has higher price elasticity than residential and commercial sector. Consumption elasticity for the season with the highest consumption level is displayed in Table 4.

Table 4. Elasticity Summary of Existing Literature

| Category | | Residential | Commercial | Industrial | Electrical |
|---|---|---|---|---|---|
| Arora (2014) | | -0.24 | | | |
| | short-term | -0.11 | -0.09 | -0.16 | -0.15 |
| | long-term | -0.20 | -0.23 | -0.57 | -0.47 |
| | short-term | -0.15 | -0.12 | -0.22 | -0.12 |
| | long-term | -0.53 | -0.53 | -1.24 | -0.80 |
| Davis and Muehlegger (2010) | | -0.28 | -0.21 | -0.71 | |
| Dahl and Roman (2004) | | | | | -0.32 |
| 2015 NEMGas Model | | -0.53 | -0.53 | -1.24 | -0.80 |
| Range | | (-0.53,-0.11) | (-0.53,-0.09) | (-1.24,-0.16) | (-0.80,-0.12) |
| Average | | -0.29 | -0.29 | -0.69 | -0.44 |
| | Short-term | -0.07 | -0.04 | -0.11 | -0.31 |
| | Long-term | -0.10 | -0.06 | -0.49 | -0.64 |

Table 5. Residential Sector Consumption Estimation Result

| Winter ( Dec. - Feb.) | | | Summer ( May. - Sep.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| (Intercept) | -5.51 | 10.46 | (Intercept) | 58.88 | 10.76 | (Intercept) | 8.91 | 16.67 |
| log(PriceWithCPI) | -0.07 | 0.03 | log(PriceWithCPI) | 0.00 | 0.03 | log(PriceWithCPI) | -0.03 | 0.04 |
| log(Residential_p) | 0.33 | 0.02 | log(Residential_p) | 0.52 | 0.02 | log(Residential_p) | 0.51 | 0.03 |
| log(Year) | 1.11 | 1.39 | log(Year) | -7.10 | 1.41 | log(Year) | -0.78 | 2.2 |
| log(HDD_Total) | 0.78 | 0.04 | Jun. | -0.16 | 0.02 | log(HDD_Total) | 0.41 | 0.03 |
| Feb. | -0.08 | 0.01 | Jul. | -0.13 | 0.02 | Apr. | -0.11 | 0.02 |
| Dec. | 0.04 | 0.01 | Aug. | -0.10 | 0.03 | Oct. | 0.28 | 0.06 |
| Region Dummy | | | Sep. | 0.04 | 0.03 | Nov. | 0.35 | 0.03 |
| | | | Region Dummy | | | Region Dummy | | |

Note: The data ranged from 1997 to 2016. log(Residential_p) represents the residential consumption level of prior month in log.

Table 6. Commercial Sector Consumption Estimation Result

| Winter ( Dec. - Feb.) | | | Summer ( May. - Sep.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| (Intercept) | -45.64 | 9.05 | (Intercept) | -21.39 | 9.33 | (Intercept) | -48.8 | 14.53 |
| log(PriceWithCPI) | -0.04 | 0.02 | log(PriceWithCPI) | -0.04 | 0.02 | log(PriceWithCPI) | -0.02 | 0.04 |
| log(Commercial_p) | 0.33 | 0.02 | log(Commercial_p) | 0.53 | 0.02 | log(Commercial_p) | 0.5 | 0.02 |
| log(HDD_Total) | 0.55 | 0.04 | log(Year) | 3.43 | 1.22 | log(HDD_Total) | 0.21 | 0.02 |
| log(Year) | 6.53 | 1.21 | Jun. | -0.06 | 0.01 | log(Year) | 6.96 | 1.91 |
| Feb. | -0.07 | 0.01 | Jul. | -0.01 | 0.01 | Apr. | -0.14 | 0.02 |
| Dec. | 0.01 | 0.01 | Aug. | 0.03 | 0.01 | Oct. | 0.11 | 0.03 |
| Region Dummy | | | Sep. | 0.08 | 0.01 | Nov. | 0.2 | 0.02 |
| | | | Region Dummy | | | Region Dummy | | |

Note: The data ranged from 1997 to 2016. log(Commercial_p) represents the commercial consumption level of prior month in log.

Table 7. Industrial Sector Consumption Estimation Result

| Winter (Nov. - Mar.) | | | Summer ( Jun. - Aug.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| (Intercept) | 3.10 | 0.32 | (Intercept) | 2.76 | 3.14 | (Intercept) | 3.03 | 0.33 |
| log(PriceWithCPI) | -0.06 | 0.01 | log(PriceWithCPI) | -0.11 | 0.17 | log(PriceWithCPI) | -0.04 | 0.01 |
| log(Industrial_p) | 0.68 | 0.03 | log(Industrial_p) | 0.77 | 0.25 | log(Industrial_p) | 0.73 | 0.03 |
| log(HDD_Total) | 0.10 | 0.02 | Jul. | 0.04 | 0.02 | log(HDD_Total) | 0.01 | 0 |
| Feb. | -0.09 | 0.01 | Aug. | 0.05 | 0.02 | May. | 0.04 | 0.01 |
| Mar. | -0.02 | 0.01 | Region Dummy | | | Sep. | 0.06 | 0.01 |
| Nov. | 0.01 | 0.01 | | | | Oct. | 0.13 | 0.01 |
| Dec. | 0.02 | 0.01 | | | | Region Dummy | -0.19 | 0.02 |
| Region Dummy | | | | | | | | |

Note: The data ranged from 2006 to 2016. log(Industrial_p) represents the industrial consumption level of prior month in log.

Table 8. Electrical Sector Consumption Estimation Result

| Winter ( Nov. - Mar.) | | | Summer ( Jun. - Aug.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| (Intercept) | 0.47 | 0.73 | (Intercept) | 3.09 | 0.47 | (Intercept) | 1.51 | 0.45 |
| log(Price_ratio) | -0.15 | 0.08 | log(Price_ratio) | -0.31 | 0.09 | log(Price_ratio) | -0.21 | 0.10 |
| log(Electrical_p) | 0.80 | 0.05 | log(Electrical_p) | 0.31 | 0.07 | log(Electrical_p) | 0.77 | 0.06 |
| log(HDD_Total) | 0.17 | 0.08 | log(CDD_Total) | 0.64 | 0.08 | log(CDD_Total) | 0.05 | 0.02 |
| Feb. | -0.11 | 0.04 | Jul. | -0.06 | 0.05 | May. | -0.01 | 0.06 |
| Mar. | 0.08 | 0.05 | Aug. | -0.11 | 0.04 | Sep. | -0.28 | 0.08 |
| Nov. | 0.00 | 0.05 | Region Dummy | | | Oct. | -0.06 | 0.04 |
| Dec. | 0.09 | 0.04 | | | | Region Dummy | | |
| Region Dummy | | | | | | | | |

Note: The data ranged from 2006 to 2016. log(Electrical_p) represents the electrical consumption level of prior month in log. Price_ratio denotes the price ratio of natural gas to coal.

## 3.2 Supply Estimation

As discussed in chapter 2, the producer's decision is made based on three factors: production cost, expected next month price and expected next year price. The specification is inspired by Deaton and Laroque (2003). They believe that in the short run, there is persistent production lag. In the long run, commodity supply is elastic and the production growth rate is mainly driven by "the excess of the current price over the long-run supply price." Marginal cost can be used to represent long-rum supply price. In this study, production is specified as following:

$$Pr^{(t,m)} = \beta_0 Pr^{(t,m-1)} \left(\frac{E_{t-1}(P_{mean}^{(t)})}{cost^{(t)}}\right)^{\beta_1} \left(\frac{E_{t,m}(p^{(t,m+1)})}{E_{t-1}(P_{mean}^{(t)})}\right)^{\beta_2} \varepsilon_{pr},$$

or equivalently, in log form:

$$ln(Pr^{(t,m)}) = \beta_0' + ln\left(Pr^{(t,m-1)}\right) + \beta_1 ln\left(\frac{E_{t-1}\left(P_{mean}^{(t)}\right)}{cost^{(t)}}\right) + \beta_2 ln\left(\frac{E_{t,m}(p^{(t,m+1)})}{E_{t-1}\left(P_{mean}^{(t)}\right)}\right) + \varepsilon_{pr}'.$$

Intuitively, production highly depends on prior production level and responds to the ratio of expected price and production cost. Term $\left(E_{t-1}\left(P_{mean}^{(t)}\right)/cost^{(t)}\right)$ represents the investment incentive: if expected price next year is higher than production cost, producers tend to increase production capacity and thus gas supply. The last term $\left(E_{t,m}(p^{(t,m+1)}/E_{t-1}(P_{mean}^{(t)})\right)$ indicates one month ahead incentive: if the expected price next month is higher than the prior expectation based on which investment decision is made, production tends to be higher. $\beta_1$ and $\beta_2$ roughly represents the investment and short-term price elasticity respectively.

The estimation is conducted at national level due to data availability. NYMEX natural gas future price is used as producers make production decisions based on expected price instead of realized price. The natural gas futures contract prices are based on Henry Hub delivery. For one month ahead expected price, we use front month contract, of which the delivery month is the calendar month following the trade date. For expected price next year, it is calculated by averaging prices of all twelve months' futures maturing next year traded at the beginning of current year. In January 2017, the expected next year price is the average of twelve futures price traded in January 2017: the futures contract maturing in January, February, March and all the months up to December 2018. The futures contract price is obtained from Quandl.

The data used to estimate natural gas production ranges from 2010 to 2016. Daily average production volume instead of monthly total production is used to account for variance in day counts. Only the most recent data is utilized to better capture the post shale gas revolution period feature and to avoid inaccuracy resulting from structural change. The production cost during 2010 ~ 2016 is set at $2.56 per MMBtu. The cost level is close to specification of the natural gas supply curve used in Deloitte's 2011 research paper and is adjusted to be consistent with observed price and production level.

The estimation result is displayed in Table 9.

Table 9. Production Estimation

| | Estimate | Std. Error | t-stats | P-value |
|---|---|---|---|---|
| $\beta_0'$ | -0.002 | 0.003 | -0.48 | 0.63 |
| $\beta_1$ | 0.014 | 0.005 | 2.53 | 0.01 |
| $\beta_2$ | 0.009 | 0.005 | 1.66 | 0.10 |

## 3.3 Storage Cost and Convenience Yield

In the non-arbitrage condition, the net marginal carry cost is defined as the storage cost net benefit of holding one unit of natural gas from period t to period (t+1). The net carry cost has several components. First, the inventory holders need to pay for physical storage. Physical storage cost is incurred for commodity storage. Secondly, there is convenience yield by holding commodity at hand. Third, the spot price might fall during the inventory period. In addition, there is opportunity cost associated with the capital invested in storage, which can be calculated as the foregone interest. This component is incorporated in the discounting of expected future price already.

The concept of convenience yield is proposed to explain the observation of price backwardation: if there is no benefit generated from stock and only physical storage cost is paid for holding commodity to the next period, price should increase indefinitely when there is stock. Kaldor (1939) and Working (1949) observed inverted inter-temporal spread for wheat market when inventory is held. Convenience yield is the benefit of holding one unit of commodity at hand which cannot be obtained by entering into futures contract. Holding inventory can help avoid stock out and production disruptions due to commodity unavailability. To a less extreme extent compared to stock out, inventory can reduce the cost of scheduling production and consumption. In addition, inventory can enable the holder to take advantage of a potential price increase.

There are many different methods to estimate convenience yield of natural gas. Hochradl and Rammerstorfer (2012) use three different methods to calculate convenience yield: traditional net convenience yield through non-arbitrage condition, look-back option based approach (sell the asset at its highest price during the time period [t, T]) and geometric Asian option based approach (trade the asset at geometric average prices during period [t, T]). The driving factors of natural gas convenience yield are also intensively discussed (Kremser and Rammerstorfer, 2010; Wei and Zhu, 2006; Volmer, 2011; Pindyck, 2001). The driving factors of convenience yield are mainly storage level, along with commodity spot prices and its variance, futures prices and its variance, and other factors such as weather and oil prices.

In this model, the net marginal carry cost incorporates both storage cost and convenience yield. Net carry cost is defined as storage cost minus convenience yield. As both storage cost and convenience yield are not easily observable, they are hard to estimate separately. This study instead estimates net marginal storage cost. Net marginal storage cost equals storage cost minus convenience yield. Net storage cost is defined as the difference between the discounted nearest month futures price and the current period spot price, as its definition illustrated.

$$sc'^{(m)} = k - cy = \frac{P_{future}}{1+r} - P_{spot}, \tag{12}$$

where r is the discount rate and $P_{futures}$ denotes the nearest month futures price.

As discussed in many papers, convenience yield is driven by expectation about future availability. It is negatively related with storage level. When there is plenty of natural gas available, the inventory holder gains little from storage as there is a minimal possibility that stock out will happen. On the contrary, when storage goes down relatively to average storage levels in a month, the market expects low availability of natural gas in the future as well and price goes up. Convenience yield goes up along with low inventory.

This study employs a monthly model to characterize monthly variation of natural gas market. This poses the difficulty of determining which level can be deemed high or low because natural gas inventory differs significantly among months. The idea of normal storage level is proposed to determine relative inventory for each month. Normal storage level will be used for convenience yield calculation. It denotes the normal storage level each month should reach to balance seasonal demand-supply relationship. If realized storage level is higher than the normal level, price goes down. If storage level is lower than normal level, price increases to reduce consumption and thus introduce more storage.

Normal storage level of each month is derived from the gap between production and total demand (including consumption demand and export) of each month, or equivalently the net withdrawal from storage each month.

$$Withdrawal^{(m)} = s^{(m-1)} - s^{(m)},$$

We take one year as the natural gas storage cycle horizon. The expected storage level of month m should be able to cover all the production-demand gaps or storage withdrawal need for the rest of the year, that is from month (m+1) to end of the year:

$$S_{norm0}^{(m)} = \sum_{i=m+1}^{12} Withdrawal^{(i)} = \sum_{i=m+1}^{12} \left(D^{(m)} + X^{(m)} - Pr^{(m-1)}\right),$$

For example, $S_{norm0}^{(Jan)}$ is the summation of the expected storage withdrawal from February to December, and $S_{norm0}^{(Nov)}$ is the expected withdrawal in December.

With the estimated net withdrawal from consumption, export and production of the year, we can estimate the trend of normal storage level as $S_{norm0}^{(m)}$. As illustrated in Figure 9, the $S_{norm0}^{(m)}$ only captures the trend of inventory level throughout the year. As the storage level cannot be negative, and in fact needs to have a minimum storage (pipeline storage) to keep the facility functioning, we re-scale by shifting $S_{norm0}^{(m)}$ up to $S_{norm}^{(m)}$ and assume a constant yearly minimum storage level at $S_{base}$. As March is always the month with lowest inventory level throughout a year, March inventory level is used to approximate $S_{base}$. The natural gas storage level in March varies in different years due to random weather or other shocks. The historical data shows that the lowest storage level of each year is close except for the years with extreme weather. In this study the base storage level $S_{base}$ is estimated by averaging the lowest working gas storage level each year (occurs in March) from 2010 to 2016 and it is 1.7 tcf. The normal storage level is thus defined in equation (13):

$$S_{norm}^{(m)} = S_{norm0}^{(m)} + \left(S_{base} - min\left(\left\{S_{norm0}^{(i)}\right\}\right)\right), i = 1,2 \ldots ,12. \tag{13}$$

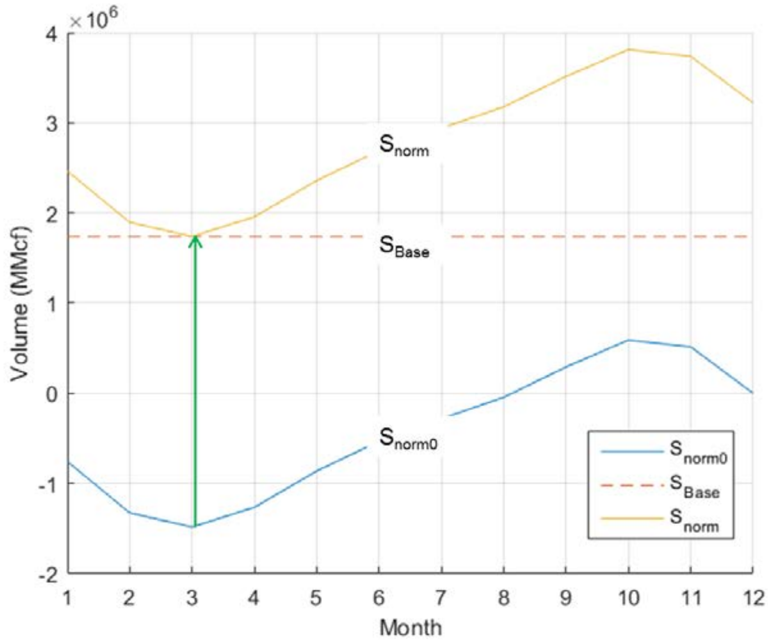


Figure 9. Illustration of normal storage level

As discussed above, the net carry cost consists of two components, the convenience yield and the physical storage cost. When the storage level is low compared to the normal storage level, the convenience yield dominates the net carry cost. The net carry cost thus drops significantly to increase storage level. Similarly, when the storage level is too high and close to reaching the storage capacity limit, the physical storage cost dominates the net carry cost. Net carry cost increases so that inventory level drops. A descriptive curve is used to show the general response of the net carry cost to the storage level. Based on the seasonal consumption and storage pattern, the net carry cost estimation is divided into four different stages that correspond to different seasons. Net marginal storage cost is defined as a function of working gas in storage, similar to Rui and Miranda (1995).

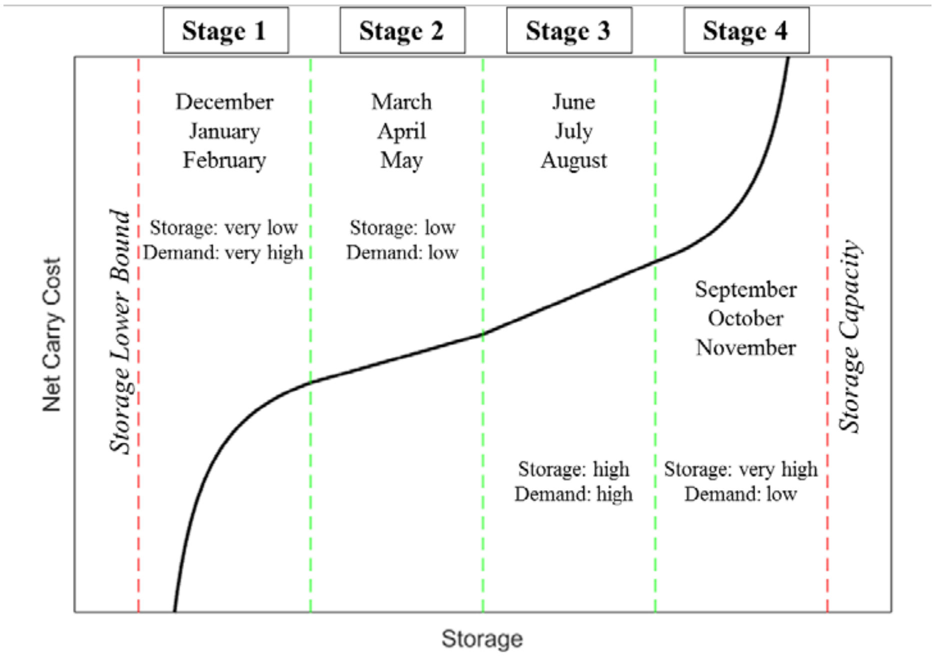


Figure 10. Descriptive illustration of net carry cost response to storage level

Stage 1: In winter season (Dec, Jan and Feb), due to the high natural gas need resulting from heating purpose, consumption is always higher than the level producers can supply, so net withdrawal from storage is necessary. As a result, the storage level of winter is low at the end of winter. In this case, when the actual storage level is lower than the normal storage level, it would cause the market to panic. To guarantee the natural gas storage to be sufficient for future usage, the convenience yield will increase rapidly and dominate the net carry cost. The spot price therefore will increase accordingly, and consequently it will reduce the consumption and hold the storage for later use. In this stage, the net carry cost response is assumed to be non-linear, and the sensitivity will be higher as the storage level moves to the lower bound (represented by $l$ in the net storage function) from the normal storage level. To represent this behavior, we use a non-linear function to fit the historical observations.

$$sc'^{(m)}\left(s^{(m)}\right) = \frac{\beta_1}{s^{(m)}/s_{norm}^{(m)} - l} + \beta_2^{(m)}, \qquad (m = Jan, Feb, Dec).$$

<u>Stage 2</u>: In spring season (Mar, Apr and May), the peak of the consumption due to heating purpose has passed. The total consumption drops from high level and the gas inventory starts to recover. The two components of the net carry cost, convenience yield and physical storage cost, both respond to the storage level. When storage level is lower than the normal storage level, the convenience yield will be higher and the physical storage cost will be lower, and vice versa. The net carry cost increases with storage level. Since the consumption level is low in spring, the market will be less responsive to the storage level as there is still time to accumulate inventory. The sensitivity of the net carry cost in general is small. We simply use a linear function to represent the correlation in this stage.

$$sc'^{(m)}\left(s^{(m)}\right) = \beta_1 \times \left(s^{(m)}/s_{norm}^{(m)}\right) + \beta_2^{(m)}, \qquad (m = Mar, Apr, May).$$

<u>Stage 3</u>: In summer season (Jun, Jul and Aug), the condition is similar to spring: the total consumption is lower than the producer can produce, and the storage will keep accumulating. Both convenience yield and physical storage cost respond to the storage level, and the net carry cost positively correlates to the storage level. The consumption need become higher and forms a small peak in summer due to higher electricity consumption resulting from cooling need. As a result, the market will react relatively stronger to the storage compared to spring. In this stage, we also use a linear function for the net carry cost-storage correlation, but expect higher sensitivity.

$$sc'^{(m)}\left(s^{(m)}\right) = \beta_1 \times \left(s^{(m)}/s_{norm}^{(m)}\right) + \beta_2^{(m)}, \qquad (m = Jun, Jul, Aug).$$

<u>Stage 4</u>: In fall season (Sep, Oct and Nov), after the storage accumulation all the way from spring to summer, the storage level is very high and close to reach the capacity limit. Moreover, the consumption in fall is generally low compared to the production level. As a result of the low consumption and sufficient supply, the convenience yield accounts for a limited

portion of the net carry cost. On the other hand, the storage level in fall is high. Most of the time it's comparable to the low-cost depleted field storage capacity, and sometimes the storage level can even exceed the depleted field capacity, especially in October. In this case, the physical storage cost of natural gas will increase rapidly and dominate the total net carry cost. To account for this behavior, instead of using the ratio of storage to normal storage level, we use the ratio of storage to the depleted field capacity for net carry cost calibration. When the ratio is approaching or exceeding 1, the net carry cost will take off and increase accordingly. As the unit cost of the other storage facilities like salt cavern or aquifers are different and much higher than the depleted field storage, a non-linear correlation is assumed with an upper bound (represented by $u$ in function below) as the total capacity of all types of storage.

$$sc'^{(m)}\left(s^{(m)}\right) = \frac{\beta_1}{s^{(m)}/s_{capacity} - u} + \beta_2^{(m)}, \qquad (m = Sep, Oct, Nov).$$

With the above four-stage division, the net carry cost function is defined as equations (1-4), and the coefficients ($\beta$'s) together with $u$ and $l$ of the functions are calibrated with historical observations. The actual storage level and depleted field storage capacity are observable from EIA database, the historical net carry cost data is calculated via equation (12) with historical prices and futures, and the normal storage level is determined from equations (13) with historical consumption and production. The calibrated functions are shown in Figure 11.

Based on the above estimated functions, we can get the storage decision, production and consumptions as endogenous functions that are derivable from price. They therefore are functions of state variables ($hdd^{(m)}$, $cdd^{(m)}$, $\varepsilon_D^{(m)}$ and $TS^{(m)}$), by substituting the price function into consumption function, production function, and non-arbitrage conditions.

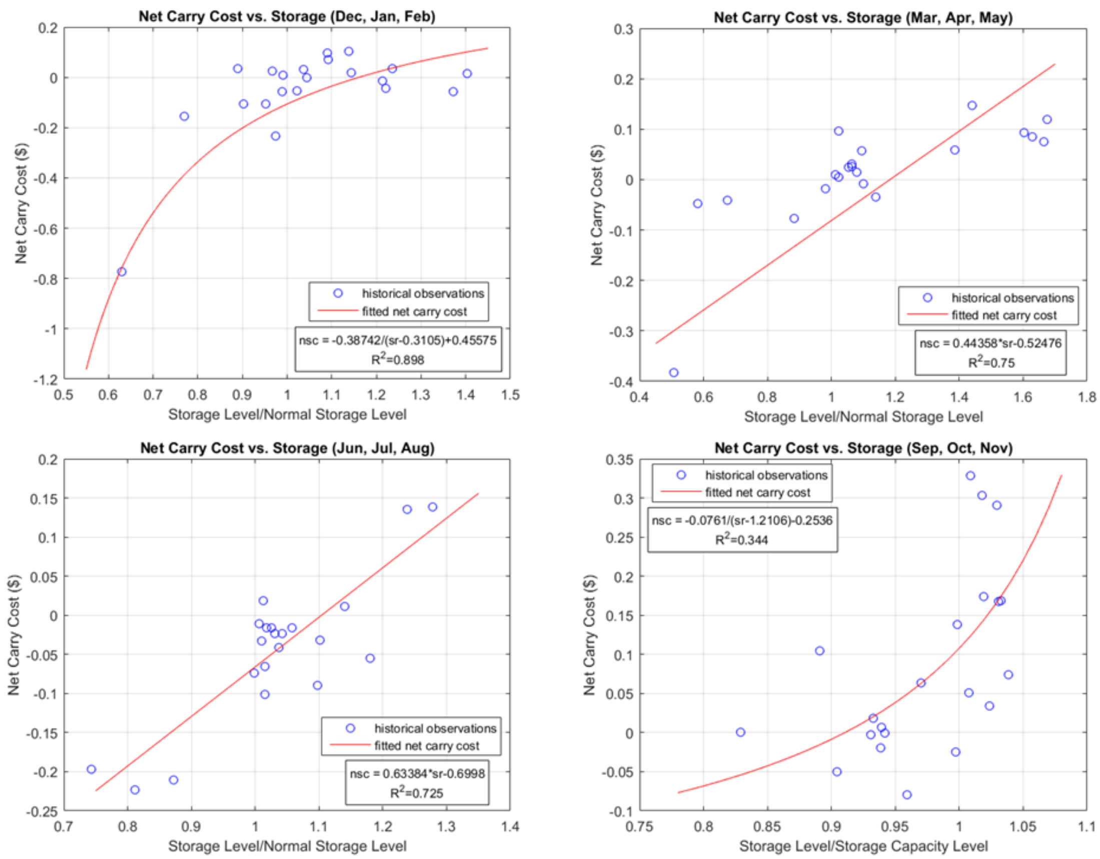


Figure 11. Net carry cost calibration results with historical observations

# CHAPTER 4. SOLVING ALGORITHM

One practical difficulty of estimating dynamic stochastic programming problems is the number of state variables included. This study tries to capture the major state variables while maintaining reasonable running time.

## 4.1 Price Function Specification

As expressed in equation (11), the current price is a function of the current weather condition, demand shock and total supply. Due to the complexity of the demand function, the explicit form of the price function is not possible to obtain in this study; therefore, a numerical approach is applied for approximation. The price function is approximated with a linear combination of independent Chebychev basis functions[3] of weather conditions in each region ($\left\{hdd_j^{(m)}, cdd_j^{(m)} | j = 1, \dots, 9\right\}$, consumption shock ($\varepsilon_D^{(m)}$), and total supply ($TS^{(m)}$) of the month. More specifically, the price is expressed as

$$\begin{aligned} p^{(m)} &= p\left(hdd^{(m)}, cdd^{(m)}, \varepsilon_D^{(m)}, TS^{(m)}\right) \approx \hat{p}\left(hdd^{(m)}, cdd^{(m)}, \varepsilon_D^{(m)}, TS^{(m)}\right) \\ &= \sum_{j=1}^{9}\left(\sum_{k_1=1}^{n_1}\sum_{k_3=1}^{n_3}\sum_{k_4=1}^{n_4} c_{jk_1k_3k_4}\phi_{j1}^{(k_1)}\left(hdd_j^{(m)}\right)\phi_3^{(k_3)}\left(\varepsilon_D^{(m)}\right)\phi_4^{(k_4)}\left(TS^{(m)}\right)\right. \\ &\left.+\sum_{k_2=1}^{n_2}\sum_{k_3=1}^{n_3}\sum_{k_4=1}^{n_4} c_{jk_2k_3k_4}\phi_{j2}^{(k_2)}\left(cdd_j^{(m)}\right)\phi_3^{(k_3)}\left(\varepsilon_D^{(m)}\right)\phi_4^{(k_4)}\left(TS^{(m)}\right)\right), \end{aligned} \tag{14}$$

where $\left\{\left(\phi_{j1}^{(1)}, \dots, \phi_{j1}^{(n_1)}\right)\right\}$ and $\left\{\left(\phi_{j2}^{(1)}, \dots, \phi_{j2}^{(n_2)}\right)\right\}$ are the univariate basis functions for $hdd_j^{(m)}$ and $cdd_j^{(m)}$ of region j; $\left\{\left(\phi_3^{(1)}, \dots, \phi_3^{(n_3)}\right)\right\}$ are the basis functions for demand shock of natural

---

[3] Chebychev basis function is of the following format: $T_0(x) = 1, T_1(x) = x, T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$ Essentially, they are polynomials.

gas; and $\left\{\left(\phi_4^{(1)}, \dots, \phi_4^{(n_3)}\right)\right\}$ are the basis functions for the total supply of natural gas. $n_1$, $n_2$, $n_3$ and $n_4$ are number of Chebychev basis function, for hdd, cdd, $\varepsilon_D$ and TS respectively. Chebychev nodes are applied in this study to solve the price function. As a result, the coefficients of equation (14), c's, is a vector of $N = 9\,(regions) \times (n_1 n_3 n_4 + n_2 n_3 n_4)$ elements. The coefficients will be solved to satisfy the competitive storage model defined above at each of the nodes.

Based on the above estimated price functions, we can solve the storage decision, production and consumption as endogenous functions that are derivable from price. They therefore are functions of the state variables ($hdd^{(m)}$, $cdd^{(m)}$, $\varepsilon_D^{(m)}$ and $TS^{(m)}$), by substituting the price function into the demand function, production function, and non-arbitrage condition.

## 4.2 Gaussian Quadrature for Integration

Substitute the net marginal storage cost function into the non-arbitrage conditions, we have

$$\frac{1}{1+r} E_m\left(p^{(m+1)}\right) = p^{(m)} + sc'\left(s^{(m)}\right). \tag{15}$$

With substitution of the price function, the uncertainty of the expected price comes from the unknown weather conditions ($hdd^{(m+1)}$ and $cdd^{(m+1)}$) at regional level, demand shock and production shock.

$$\begin{aligned} E_m\left(p^{(m+1)}\right) &= E_m\left(\hat{p}^{(m+1)}\left(hdd^{(m+1)}, cdd^{(m+1)}, \varepsilon_D^{(m+1)}, TS^{(m+1)}\right)\right) \\ &= E_m\left(\hat{p}^{(m+1)}\left(hdd^{(m+1)}, cdd^{(m+1)}, \varepsilon_D^{(m+1)}, TP^{(m)}\varepsilon_{Pr}^{(m)} + s^{(m)}\right)\right) \\ &= E_m\left(\hat{p}^{(m+1)}\left(hdd^{(m+1)}, cdd^{(m+1)}, \varepsilon_D^{(m+1)}, TP^{(m)}\left(Pr^{(m-1)}, E_m\left(p^{(m+1)}\right)\right)\varepsilon_{Pr}^{(m)} + s^{(m)}\right)\right), \end{aligned} \tag{16}$$

As the total production $TP^{(m)}$, is a function of the expected price, equation (16) can be expressed as an equation of expected price, storage level ($s^{(m)}$) and state variables. Thus, the expected price can be solved as a function of state variables and storage level.

The equilibrium condition can be re-written as below by substitution of price function and consumption function.

$$TS^{(m)} = D_{tot}^{(m)} + s^{(m)} = \sum_{i=1}^{4}\sum_{j=1}^{9} D_{ij}\left(D^{(m-1)}, p^{(m)}, hdd^{(m)}, cdd^{(m)}\right)\varepsilon_D^{(m)} + s^{(m)}, \tag{17}$$

With the equations 15-17, the equation set is closed and ready to be solved for price function and storage level with the method described later.

To get the expected price, we need to integrate price over weather, demand shock and production shock distributions. Weather distribution varies among regions and more importantly correlates with one region to another. It is difficult to come up with a parametric distribution fitting regional level weather while maintaining the spatial correlations. To be consistent with reality and maintain the spatial correlations, empirical distribution of regional level weather is used in this study, which is observed from historical data of the past 28 years (1989 – 2016).

Both consumption and production shock are random variables following Gaussian distributions with mean μ = 1 and their standard deviation specified to be consistent with historical observation. Consumption shock $\varepsilon_D^{(m)}$ is calculated from the ratio of historical consumption to the consumption level predicted using estimated consumption function. Production shock $\varepsilon_{Pr}^{(m)}$ is calculated in the same way. The standard deviation of both consumption and production shock is computed from the historical estimated number and used to define their distribution. The Gaussian quadrature is applied to calculate the expectation as weighted average over Gaussian nodes, and in this study, 25 = 5×5 mesh grids of Gaussian nodes are used for consumption shock and production shock. Therefore, the expected price can be calculated as the integration over a total number of 700 = 28×25 grids of historical weather realizations, consumption shock and production shock:

$$
\begin{aligned}
&E_m\left(p^{(m+1)}\right)) = E_m\left(\hat{p}^{(m+1)}\left(hdd^{(m+1)}, cdd^{(m+1)}, \varepsilon_D^{(m+1)}, Pr^{(m)}\left(E_m\left(p^{(m+1)}\right)\right)\varepsilon_{Pr}^{(m)} + s^{(m)}\right)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\iint \hat{p}^{(m+1)}\left(hdd_i^{(m+1)}, cdd_i^{(m+1)}, \varepsilon_D^{(m+1)}, Pr\left(Pr^{(m-1)}, E_m\left(p^{(m+1)}\right)\right)\varepsilon_{Pr}^{(m)} + s^{(m)}\right) d\varepsilon_D d\varepsilon_{Pr} \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{5} w_{jk}\hat{p}^{(m+1)}\left(hdd_i^{(m+1)}, cdd_i^{(m+1)}, \varepsilon_{D,j}^{(m+1)}, Pr\left(Pr^{(m-1)}, E_m\left(p^{(m+1)}\right)\right)\varepsilon_{Pr,k}^{(m)} + s^{(m)}\right).
\end{aligned}
\tag{18}
$$

The weather variables (HDD and CDD) for current month and next month are assumed to be independent at regional level. That is, current month weather conditions provide no information regarding what the next month weather will be. The independency assumption is supported by historical data. Table 10 lists the regional level monthly weather dependency, using HDD number. Using CDD yields the same result as the both of them are derived from temperature. A2M1 represents HDD of climate region 2 in January. Taking representative winter month – January for example, the AR(1) coefficient for HDD ranged from 0.2 to 0.3, small in general. Out of the total nine climate regions, only one region has AR(1) coefficient significant at 5% level in January. The HDDs of other winter seasons like February, March, October, November and December yield similar results. For all nine climate regions in six winter months, only 8 out of 54 have AR(1) coefficient significant at 5% level. The assumption of time independence of weather variables may not hold strictly but making the assumption greatly facilitates solving the problem.

Table 10. HDD Monthly Dependency Check

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A1M1 | A1M2 | A1M3 | A1M4 | A1M5 | A1M6 | A1M7 | A1M8 | A1M9 | A1M10 | A1M11 | A1M12 |
| AR(1) | 0.28 | 0.23 | 0.01 | 0.31 | -0.02 | 0.01 | 0.91 | -0.40 | 0.48 | 0.13 | 0.47 | 0.03 |
| P-value | 0.12 | 0.21 | 0.92 | 0.07 | 0.63 | 0.79 | 0.19 | 0.67 | 0.13 | 0.72 | 0.08 | 0.88 |
| | A2M1 | A2M2 | A2M3 | A2M4 | A2M5 | A2M6 | A2M7 | A2M8 | A2M9 | A2M10 | A2M11 | A2M12 |
| AR(1) | 0.29 | 0.38 | 0.19 | 0.32 | -0.07 | 0.14 | 1.22 | -0.58 | 0.34 | 0.19 | 0.48 | 0.12 |
| P-value | 0.12 | 0.03 | 0.07 | 0.07 | 0.29 | 0.10 | 0.00 | 0.27 | 0.17 | 0.59 | 0.04 | 0.53 |
| | A3M1 | A3M2 | A3M3 | A3M4 | A3M5 | A3M6 | A3M7 | A3M8 | A3M9 | A3M10 | A3M11 | A3M12 |
| AR(1) | 0.33 | 0.51 | 0.09 | 0.27 | -0.05 | 0.04 | 0.42 | 1.55 | 0.15 | -0.26 | 0.71 | 0.08 |
| P-value | 0.04 | 0.00 | 0.34 | 0.23 | 0.31 | 0.34 | 0.20 | 0.18 | 0.68 | 0.34 | 0.03 | 0.69 |
| | A4M1 | A4M2 | A4M3 | A4M4 | A4M5 | A4M6 | A4M7 | A4M8 | A4M9 | A4M10 | A4M11 | A4M12 |
| AR(1) | 0.24 | 0.29 | 0.52 | 0.50 | 0.49 | 0.12 | 0.43 | 0.38 | -0.18 | -0.49 | -0.13 | 0.30 |
| P-value | 0.26 | 0.07 | 0.01 | 0.01 | 0.00 | 0.10 | 0.01 | 0.26 | 0.49 | 0.07 | 0.50 | 0.16 |
| | A5M1 | A5M2 | A5M3 | A5M4 | A5M5 | A5M6 | A5M7 | A5M8 | A5M9 | A5M10 | A5M11 | A5M12 |
| AR(1) | 0.32 | 0.21 | 0.07 | 0.06 | -0.03 | -0.02 | 0.37 | -5.09 | 0.29 | 0.31 | 0.17 | -0.04 |
| P-value | 0.11 | 0.22 | 0.47 | 0.35 | 0.03 | 0.32 | 0.44 | 0.50 | 0.69 | 0.50 | 0.45 | 0.82 |
| | A6M1 | A6M2 | A6M3 | A6M4 | A6M5 | A6M6 | A6M7 | A6M8 | A6M9 | A6M10 | A6M11 | A6M12 |
| AR(1) | 0.18 | 0.18 | 0.00 | 0.16 | 0.07 | 0.01 | 0.01 | 5.32 | 1.72 | -0.01 | 0.38 | 0.09 |
| P-value | 0.25 | 0.43 | 0.98 | 0.17 | 0.08 | 0.55 | 0.96 | 0.35 | 0.01 | 0.98 | 0.25 | 0.55 |
| | A7M1 | A7M2 | A7M3 | A7M4 | A7M5 | A7M6 | A7M7 | A7M8 | A7M9 | A7M10 | A7M11 | A7M12 |
| AR(1) | 0.28 | 0.03 | 0.11 | 0.33 | 0.11 | 0.06 | 0.92 | 0.38 | 0.76 | 0.12 | 0.05 | 0.14 |
| P-value | 0.11 | 0.84 | 0.55 | 0.02 | 0.35 | 0.11 | 0.00 | 0.64 | 0.02 | 0.70 | 0.77 | 0.55 |
| | A8M1 | A8M2 | A8M3 | A8M4 | A8M5 | A8M6 | A8M7 | A8M8 | A8M9 | A8M10 | A8M11 | A8M12 |
| AR(1) | 0.36 | -0.03 | 0.38 | 0.34 | 0.18 | 0.02 | 0.03 | -0.20 | 2.13 | 0.45 | 0.25 | 0.36 |
| P-value | 0.06 | 0.88 | 0.01 | 0.04 | 0.00 | 0.13 | 0.89 | 0.88 | 0.00 | 0.18 | 0.10 | 0.12 |
| | A9M1 | A9M2 | A9M3 | A9M4 | A9M5 | A9M6 | A9M7 | A9M8 | A9M9 | A9M10 | A9M11 | A9M12 |
| AR(1) | 0.18 | 0.15 | 0.23 | 0.24 | -0.04 | 0.13 | 0.78 | 0.30 | 0.08 | -0.08 | 0.10 | 0.20 |
| P-value | 0.41 | 0.38 | 0.03 | 0.09 | 0.69 | 0.19 | 0.00 | 0.57 | 0.83 | 0.81 | 0.62 | 0.32 |

Note: A2M3 represent climate region 2, March. AR(1) denotes the AR(1) coefficient. P-value is the p-value for AR(1) coefficient.

## 4.3 Equation Sets and Iteration Method

As discussed above, equations 15 – 18 need to be solved to determine the price function and the functions that depend on price with the Gaussian quadrature method for integration approximation.

$$\frac{1}{1+r}E_m\left(p^{(m+1)}\right) = p^{(m)} + sc'(s^{(m)}) \tag{19a}$$

$$E_m\left(p^{(m+1)}\right)) \\ = \frac{1}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{5} w_{jk}\hat{p}^{(m+1)}\left(hdd_i^{(m+1)}, cdd_i^{(m+1)}, \varepsilon_{D,j}^{(m+1)}, Pr\left(Pr^{(m-1)}, E_m\left(p^{(m+1)}\right), P_{mean}^{(t)}, cost^{(t)}\right)\varepsilon_{Pr,k}^{(m)} + s^{(m)}\right) \tag{19b}$$

$$TS^{(m)} = \sum_{i=1}^{4}\sum_{j=1}^{9} D_{ij}^{m}\left(D^{(m-1)}, p^{(m)}, hdd_j^{(m)}, cdd_j^{(m)}\right)\varepsilon_D^{(m)} + s^{(m)}. \tag{19c}$$

From above equation set, when the regional consumption functions ($D_{ij}$), the total production function ($Pr$), and the net marginal storage cost function ($sc'$) format are estimated, and the production cost ($cost^{(t)}$), the prior consumption ($D^{(m-1)}$) and prior production ($Pr^{(m-1)}$) are known, the equations (19a-c) will be functions only of price ($p^{(m)}$), expected price $E_m\left(p^{(m+1)}\right)$ and storage level ($s^{(m)}$), besides the state variables ($hdd^{(m)}$, $cdd^{(m)}$, $\varepsilon_D^{(m)}$ and $TS^{(m)}$). As a result, when the coefficients of the estimated price function are given or determined, the equations are solvable.

As the demand functions vary across months, the price functions of each month are not constant across months. In fact, the price function varies across years for the same month as well because of different prior consumption and production levels. That is, the price function for January 2010 and January 2011 are different. When solving equations (19a-c), although the format of the equations is the same year over year resulting from our specification, the parameters of the equations will be different when the prior values ($D^{(m-1)}$ and $Pr^{(m-1)}$) or weather distributions are different. Therefore, the coefficients (c's) of the price function will be different for various years and months if $D^{(m-1)}$ and $Pr^{(m-1)}$ are different. Since the weather distribution is different for each month throughout the year, the price function is solved on a yearly basis when solving for multiple years. Based on this, to determine the coefficients of each month and each year, a numerical approach with iteration method is applied.

As discussed above, when a multi-year simulation is required, the price function will be solved year by year, i.e. every 12-month period is solved together using iteration algorithm and backward induction method.

Before the iteration begins, an initial guess of the yearly price mean ($P_{mean,0}^{(t)}$) and the coefficients of price functions for 12 month ($c_0^{(1)} - c_0^{(12)}$) are estimated. The iteration then begins by solving the coefficients of the first month $c_1^{(1)}$ (for example January), with the known prior values $D^{(m-1)}$ and $Pr^{(m-1)}$ from the previous December and the initial guess of February coefficients $c_0^{(2)}$. Chebyshev polynomials and nodes are applied to update the January coefficients with the iteration method described below.

Once the first month coefficient $c_1^{(1)}$ is updated, it is used to estimate the expected January consumption and production, $D^{(1)}$ and $Pr^{(1)}$, by substituting the average $hdd^{(1)}$ and $cdd^{(1)}$ into consumption and production functions[4]. These two estimations will be applied as the prior values for February, and $c_1^{(2)}$ is updated with the initial guess of March coefficients $c_0^{(3)}$. This procedure is repeated for the remaining ten months with prior value estimation and coefficients update. At the end of the 12-month iteration, the December coefficients $c_1^{(12)}$ is updated with $c_0^{(1)}$. [5]

---

[4]The expected consumption and production of current month are used as the prior values of next month, without consumption or production shock.

[5] $c_1^{(12)}$should be updated with the coefficient $c_0^{(1)}$ of next January. In this study the coefficients of January in adjacent years are assumed to be the same when solving for a given year's coefficients.

Once the 12-month coefficients ($c_1^{(1)}$-$c_1^{(12)}$) are updated, the monthly estimated price and consumption can be calculated via the price function and consumption function with the mean HDD, CDD, therefore the updated yearly price mean ($P_{mean,1}^{(t)}$) is calculated as the consumption weighted average of monthly price:

$$P_{mean}^{(t)} = \frac{\sum_{m=1}^{12} D^{(m)}\left(p^{(m)}, hdd_{mean}^{(m)}, cdd_{mean}^{(m)}\right)p^{(m)}\left(hdd_{mean}^{(m)}, cdd_{mean}^{(m)}\right)}{\sum_{m=1}^{12} D^{(m)}\left(p^{(m)}, hdd_{mean}^{(m)}, cdd_{mean}^{(m)}\right)}.$$

The updated 12-month coefficients ($c_1^{(1)}$-$c_1^{(12)}$) and yearly price mean ($P_{mean,1}^{(t)}$) are compared with the prior guess $c_0^{(1)}$-$c_0^{(12)}$ and $P_{mean,0}^{(t)}$ to the predetermined convergence level. If the difference is below the convergence level, the new coefficients are part of the solution and iteration stops for this year. Otherwise, the iteration continues to update the coefficients to replace the older guess until the convergence level is met.

When the iteration of the first year (all 12 months) is completed, the price function coefficients are determined, and the price, consumption, production, and storage level of each month can be estimated. Therefore, the prior values of the second year can be taken from the December values of the first year, and the iteration method discussed above is repeated for the second year until convergence and then the second year coefficients are determined. Repeat this procedure with the rest of years, and the price functions of all years are determined.

Iteration Method for Coefficient Update

Step 1: Define Chebychev polynomials and nodes

In this study, $n_1 = n_2 = n_3 = n_4 = 3$ independent Chebychev polynomials of each state variable ($hdd_j^{(m)}$, $cdd_j^{(m)}$, $\varepsilon_D^{(m)}$ and $TS^{(m)}$) are used approximate the price function. Thus, as discussed in equation (18), a total number of $N = 9(regions) \times (n_1 n_3 n_4 + n_2 n_3 n_4) = 486$ independent Chebychev polynomials will be constructed as

$$\left\{\phi_{j1}^{(k_1)}\phi_3^{(k_3)}\phi_4^{(k_4)}, \phi_{j2}^{(k_2)}\phi_3^{(k_3)}\phi_4^{(k_4)}\right\}, \qquad j = 1, \dots, 9; k_1, k_2, k_3, k_4 = 1, \dots, 3. \tag{20}$$

Construct the Chebychev nodes for each state variable ($hdd_j^{(m)}$, $cdd_j^{(m)}$, $\varepsilon_D^{(m)}$ and $TS^{(m)}$) of the given month *m* and each region ($j = 1, \dots, 9$) respectively. The nodes are selected over the domains, $\left[hdd_{j,min}^{(m)}, hdd_{j,max}^{(m)}\right]$, $\left[cdd_{j,min}^{(m)}, cdd_{j,max}^{(m)}\right]$, $\left[\varepsilon_{D,min}^{(m)}, \varepsilon_{D,max}^{(m)}\right]$ and $\left[TS_{min}^{(m)}, TS_{max}^{(m)}\right]$, which are defined based on the historical data of the month of each region. According to the Chebychev polynomial structure above, 3 Chebychev nodes of $\varepsilon_D^{(m)}$ and $TS^{(m)}$ respectively, and 54 nodes of $\left(hdd_j^{(m)}, cdd_j^{(m)}\right)_{j=1,\dots,9}$ are selected, thus a grid of $N = 486$ interpolation nodes are constructed by Cartesian product of weather condition nodes and total supply nodes as

$$\left\{\left(hdd_{1,k_1}^{(m)}, \dots, hdd_{9,k_1}^{(m)}, cdd_{1,k_1}^{(m)}, \dots, cdd_{9,k_1}^{(m)}, \varepsilon_{D,k_2}^{(m)}, TS_{k_3}^{(m)}\right) \middle| k_1 = 1, \dots, 54; k_2, k_3 = 1,2,3\right\}. \tag{21}$$

Since the Chebychev polynomials are all defined on the domain of $[-1,1]$, the state variable nodes need to be normalized as

$$z_k = \frac{2x_k - x_{max} - x_{min}}{x_{max} - x_{min}}, \qquad x_k \in \left\{hdd_j^{(m)}, cdd_j^{(m)}, \varepsilon_D^{(m)}, TS^{(m)}\right\}_{j=1,\dots,9}. \tag{22}$$

Step 2: Solve equation set with given coefficients of month (m+1)

Start with an initial guess of coefficients with month (m+1), $c_0^{(m+1)}$, substitute Chebychev polynomials into the price function, and apply $c_0^{(m+1)}$, we would have

$$\hat{p}^{(m+1)} = \sum_{j=1}^{9}\left(\sum_{k_1=1}^{n_1}\sum_{k_3=1}^{n_3}\sum_{k_4=1}^{n_4} c_{jk_1k_3k_4}^{(m+1)}\phi_{j1}^{(k_1)}\phi_3^{(k_3)}\phi_4^{(k_4)} + \sum_{k_2=1}^{n_2}\sum_{k_3=1}^{n_3}\sum_{k_4=1}^{n_4} c_{jk_2k_3k_4}^{(m+1)}\phi_{j2}^{(k_2)}\phi_3^{(k_3)}\phi_4^{(k_4)}\right), \tag{23}$$

Substitute equation (23) into the equation (19b), and apply the first Chebychev node $(hdd_{1,1}^{(m)}, \dots, hdd_{9,1}^{(m)}, cdd_{1,1}^{(m)}, \dots, cdd_{9,1}^{(m)}, \varepsilon_{D,1}^{(m)}, TS_1^{(m)})$. Solve equation set (19a – c) for price, expected price and storage level, denoted as $p_1^{(m)}$, $E_1^{(m)}$ and $s_1^{(m)}$. Repeat with all remaining Chebychev nodes and we will get the completed set of solutions, $\boldsymbol{p}^{(m)} = \begin{pmatrix} p_1^{(m)} \\ \vdots \\ p_N^{(m)} \end{pmatrix}$, $\boldsymbol{E}^{(m)} = \begin{pmatrix} E_1^{(m)} \\ \vdots \\ E_N^{(m)} \end{pmatrix}$ and $\boldsymbol{s}^{(m)} = \begin{pmatrix} s_1^{(m)} \\ \vdots \\ s_N^{(m)} \end{pmatrix}$.

Step 3: Update coefficients for month (m)

The $N \times N$ operational matrix $\Phi$ is defined by evaluating each of the $N$ independent Chebychev polynomials at each of the $N$ interpolation nodes.

$$\Phi_{N\times N} = \begin{bmatrix} \{\phi_{11}^{(1)}\phi_3^{(1)}\phi_4^{(1)}\}\big|_1, \dots, \{\phi_{11}^{(n)}\phi_3^{(n)}\phi_4^{(n)}\}\big|_1, \{\phi_{12}^{(1)}\phi_3^{(1)}\phi_4^{(1)}\}\big|_1, \dots, \{\phi_{12}^{(n)}\phi_3^{(n)}\phi_4^{(n)}\}\big|_1, \dots, (all\ 9\ regions), \dots, \{\phi_{91}^{(1)}\phi_3^{(1)}\phi_4^{(1)}\}\big|_1, \dots, \{\phi_{92}^{(n)}\phi_3^{(n)}\phi_4^{(n)}\}\big|_1 \\ \vdots \qquad \vdots \qquad \vdots \\ (all\ N\ (hdd, cdd, \varepsilon_D, TS)\ nodes) \\ \vdots \qquad \vdots \qquad \vdots \\ \{\phi_{11}^{(1)}\phi_3^{(1)}\phi_4^{(1)}\}\big|_N, \dots, \{\phi_{11}^{(n)}\phi_3^{(n)}\phi_4^{(n)}\}\big|_N, \{\phi_{12}^{(1)}\phi_3^{(1)}\phi_4^{(1)}\}\big|_N, \dots, \{\phi_{12}^{(n)}\phi_3^{(n)}\phi_4^{(n)}\}\big|_N, \dots, (all\ 9\ regions), \dots, \{\phi_{91}^{(1)}\phi_3^{(1)}\phi_4^{(1)}\}\big|_N, \dots, \{\phi_{9,2}^{(n)}\phi_3^{(n)}\phi_4^{(n)}\}\big|_N \end{bmatrix},$$

Therefore, we have the relation based on the price function that

$$\boldsymbol{p}^{(m)} = \begin{pmatrix} p_1^{(m)} \\ \vdots \\ p_N^{(m)} \end{pmatrix} = \Phi_{N\times N}^{(m)} \begin{pmatrix} c_{1,1}^{(m)} \\ \vdots \\ c_{1,N}^{(m)} \end{pmatrix} = \Phi \times c_1^{(m)}, \tag{24}$$

The coefficients of month (m), $c_1^{(m)}$, is updated from equation (24) as $c_1^{(m)} = \left(\Phi_{N\times N}^{(m)}\right)^{-1} \boldsymbol{p}^{(m)}$.

# CHAPTER 5. MODEL RESULTS AND DISCUSSION

This research aims to construct a competitive storage model to explain the price behavior of natural gas after the shale gas boom. With the model constructed and solved, how can we demonstrate that the model is valid and is appropriate to be used for policy analysis? If fed with historical inputs including weather variable, historical weather shocks, will the model be capable of generating price series that are consistent with historical prices? Numerically, how accurate are the price and storage policy function? The model results can be deemed as satisfactory if the solution passes accuracy tests with reasonable error range and can generate price series consistent with historical prices.

This chapter conducts accuracy tests and back-testing to validate the model solution. The implications of the model results are also analyzed.

## 5.1 Accuracy Tests

The competitive storage model is solved with the algorithm discussed in Chapter 4. Before it is applied for future simulation, we need to verify that the solution to the model is accurate enough and the assumptions made to the model are valid. The solution inaccuracy is primarily from three sources of error: (1) the approximation error from the polynomial interpolation; (2) the simulation error from using mean weather variables HDD/CDD without shock to approximate the prior values to feed into next month function; and (3) the simulation error from assuming the coefficients of price function of the same month in adjacent years to be the same. As discussed in Chapter 4, these assumptions are made to close the function sets and to achieve reasonable computation time. The impact of these assumptions is investigated respectively for model validation purpose.

Approximation error of polynomial interpolation

According to the interpolation principles (Miranda and Fackler, 2002), we choose to use Chebychev polynomial interpolation with Chebychev nodes in this study for the price approximation. With this method, the price solutions at the state variable nodes will be accurate, while solutions between nodes will yield an approximation error. As the number of nodes increase, the approximation error will be reduced.

The approximation error is developed from the following equation:

$$p_i = \hat{p}_i + e_i, \tag{25}$$

$$p(hdd_i, cdd_i, \varepsilon_{Di}, TS_i) = e_i + \sum_{j=1}^{N} c_j \phi_j (hdd_i, cdd_i, \varepsilon_{Di}, TS_i). \tag{26}$$

The left hand side of equation (26) is the actual price solution of the equation sets (19a-c) with a given set of state variables, the right hand side is the price approximation calculated with the approximated polynomial interpolation under the same condition, and $e_i$ is the approximation error of the model. In this thesis, the approximation error will be presented in a unit free form as the percentage error, $\varepsilon_i = (p_i - \hat{p}_i)/p_i$.

The accuracy test for this model is conducted as following: Starting with the prior consumption and production values from December 2009, solve the model for 7 years' horizon up to 2016 and observe the approximated price functions for each month of the 7 years. Generate state variables $(hdd_i, cdd_i, \varepsilon_{Di}, TS_i)$ in the state space for each month, with each set of state variables, the price solution $(p_i)$ is calculated by solving the equation sets (19a-c), and the approximated price is calculated directly with the price function. The percentage errors $(\varepsilon_i)$ are collected with all state variable sets of all month and all years.

100 sets of the state variables are randomly selected for each month of each year, i.e. in total 8400 cases over the 7-year period. Table 11 below shows both the maximum errors and mean errors for all the 8400 cases. To interpret the error, $\varepsilon_i = -0.1$ means that the market agent makes a prediction error of 1 dollar when they spend 10 dollars if the solved policy function is used. While $\varepsilon_i = -0.001$ means consumers will need to spend 1000 dollars to induce a 1-dollar mistake.

Table 11. Accuracy Test Results with Polynomial Approximation by Month

| Month | Jan | Feb | Mar | Apr | May | Jun |
|---|---|---|---|---|---|---|
| Max Abs. Error | 4.87E-03 | 6.38E-03 | 1.98E-03 | 1.53E-03 | 1.07E-03 | 3.79E-04 |
| Mean Abs. Error | 7.94E-04 | 1.11E-03 | 3.53E-04 | 3.69E-04 | 2.52E-04 | 1.00E-04 |
| Month | Jul | Aug | Sep | Oct | Nov | Dec |
| Max Abs. Error | 4.54E-04 | 1.02E-03 | 1.89E-03 | 5.71E-03 | 1.45E-03 | 4.17E-03 |
| Mean Abs. Error | 9.87E-05 | 1.72E-04 | 2.91E-04 | 1.76E-03 | 2.37E-04 | 7.38E-04 |

The overall average approximation error of all years and all months is 1.26e-4, which means that every 8000 dollars spent on natural gas consumption will lead to \$1 mistake. The overall maximum approximation error is 6.38e-3, and this error occurred in February. Moreover, from Table 11, we can observe that winter months (December, January, February) and October are the months with higher approximation errors. Winter months have greater variances with HDD, and the natural gas total availability is smaller in winter due to the higher consumption level. These will lead to higher sensitivity in terms of price response. October has the highest storage level of the year and usually close to the storage capacity. The price is sensitive when the total availability changes. Therefore, these months would have higher percentage error when using second order polynomial to approximate the price solution.

On the other hand, though we will get a maximum error of 0.64% (0.01% on average) from the approximation, it is far smaller than the price volatility from the historical observation (28%). Three nodes of each state variable (HDD, CDD, consumption shock, and total supply) are used in this thesis considering the balance between the computation time and solution accuracy. When using three nodes of all variables, a total number of 486 meshed grids are created for computation. If node number per variable is increased from 3 to 4 or 5, the number of meshed grids will increase to 1152 or 2250 accordingly. Test simulations are conducted with 3 and 5 nodes of each state variable for a time period of 12 months. The simulation time increased by 8 times, but the simulation results are not impacted materially. Therefore, 3 nodes of each state variable is applied in this study to save the computation time at an acceptable accuracy level.

Simulation Error from Using Mean HDD/CDD without Shock in the Model

Because an AR(1) function is used, prior consumption and production values are required in the model equation sets. To solve the equation for each month, the prior values need to be estimated from the solution of the previous month. Natural gas consumption consists of four sectors of nine regions, a mesh grid of $2^{36}$ cases will need to be generated even if each component takes only 2 nodes. This is computationally expensive to achieve. Alternatively, we take the expected consumption and production of current month as the prior values of next month, with mean HDD/CDD and no consumption shock applied. This section is to calculate the simulation error raised from this simplification.

The key concern of this assumption is the error induced by using the expected mean consumption and production as the prior values for next month to solve the model. We conduct the accuracy test by introducing some random error to the consumption and production when it is applied as the prior value of next month to solve the model. In this manner, the final solution of the price function is also changed. The simulation error is then defined as the price difference between the price simulation with and without the consumption/production shifts. The accuracy is tested with a 12-month period of the year 2016 with the consumption shock randomly selected from -3% to 3%. The simulation error results are summarized in Table 12.

Table 12. Accuracy Test Results with Prior Value Assumption

| Month | Jan | Feb | Mar | Apr | May | Jun |
|---|---|---|---|---|---|---|
| Max Abs. Error | 7.25E-03 | 5.81E-03 | 6.11E-03 | 4.96E-03 | 4.35E-03 | 3.39E-03 |
| Mean Abs. Error | 3.17E-03 | 2.65E-03 | 2.75E-03 | 2.26E-03 | 2.18E-03 | 1.45E-03 |
| Month | Jul | Aug | Sep | Oct | Nov | Dec |
| Max Abs. Error | 3.11E-03 | 4.66E-03 | 6.30E-03 | 7.60E-03 | 9.24E-03 | 6.98E-03 |
| Mean Abs. Error | 1.62E-03 | 1.99E-03 | 3.57E-03 | 3.95E-03 | 4.25E-03 | 3.34E-03 |

The overall average of the yearly inconsistency is 0.28% of all years, and the maximum inconsistency is 0.92% from the first year of simulation. Compared to price standard deviation from the historical observation (28%), the simulation error with prior value assumption is small and deemed acceptable.

<u>Simulation error from the same January price approximation in adjacent years</u>

The competitive storage model is solved with backward induction method. As discussed in Chapter 4, the coefficients of price approximation function of January ($c_t^{(1)}$) are determined by that of February ($c_t^{(2)}$), and the coefficients of December ($c_t^{(12)}$) should be determined by the coefficients of next January ($c_{t+1}^{(1)}$). To close the function sets (19a-c) and make them solvable, we have to make estimation for $c_{t+1}^{(1)}$ for the last month of each year (December). In this study, we make the assumption that the coefficients of the price approximation function is the same for the first month in adjacent years. The simulation error induced by this assumption is tested in this section.

To conduct the accuracy test, we solve the model for a 7-year time period with the starting value of December 2009, that is the model for 2010 – 2016 is solved. To validate the assumption that the price approximation is similar in the adjacent January, we calculate the price at the same randomly selected points, and compare the prices between using $c_t^{(1)}$ of 2010 and 2011. The simulation error is defined as the difference between prices calculated with $c_t^{(1)}$ of 2010 and 2011. The same process will be applied to each of all pairs of adjacent years, 2010/2011, 2011/2012, till 2015/2016. The simulation errors are collected and summarized in Table 13.

Table 13. Accuracy Test Results with Adjacent Years Assumption

| Comparison Years | 2010/11 | 2011/12 | 2012/13 | 2013/14 | 2014/15 | 2015/16 |
|---|---|---|---|---|---|---|
| Max abs. Error at Nodes | 5.05E-03 | 1.38E-03 | 2.49E-03 | 3.92E-03 | 1.45E-03 | 2.95E-03 |
| Mean abs. Error at Nodes | 3.51E-03 | 4.04E-04 | 1.00E-03 | 2.59E-03 | 4.91E-04 | 1.61E-03 |
| Max abs. Error at Non-Nodes | 9.65E-03 | 2.82E-03 | 4.75E-03 | 7.38E-03 | 2.41E-03 | 5.43E-03 |
| Mean abs. Error at Non-Nodes | 6.84E-03 | 6.20E-04 | 2.07E-03 | 5.15E-03 | 5.75E-04 | 2.97E-03 |

The overall average of the yearly inconsistency is 0.19% of all years, and the maximum inconsistency is 0.97% from the first year of simulation. Compared to the price volatility of January (36.9%) from historical observation, the simulation error is negligible.

## 5.2 Back-testing Result

With the competitive storage model set up in Chapter 3 and solving algorithm discussed in Chapter 4, a test simulation is constructed with historical observations from 2010 to 2016 to validate the model by verifying the price average, variance, and autocorrelation.

The back-testing is conducted in steps below:

1. A set of monthly pricing functions for the 7 year testing horizon is solved: the test simulation is set starting with the prior consumption/production values from December 2009, and the price functions are solved on a yearly basis for the following 7 years from 2010 to 2016. The solving algorithm is described in Chapter 4.

2. Calculate price based on observed weather variables: after the price functions are solved in step 1, given the starting values ($D_{2009}^{12}$, $s_{2009}^{12}$, and $Pr_{2009}^{12}$), the total availability of January 2010 is equal to ($s_{2009}^{12} + Pr_{2009}^{12}$). The simulated 2010 January price ($p_{2010}^{1}$) is calculated by substituting the total availability, the actual weather conditions (HDD/CDD) and consumption shock of January 2010 to the solved price function.
3. Solve for consumption and other variables following price: with the simulated price ($p_{2010}^{1}$), the simulated consumption ($D_{2010}^{1}$) can be calculated through the consumption functions for each sector and total consumption level. The simulated storage level ($s_{2010}^{1}$), expected price ($E(P_{2010}^{2})$), and production ($Pr_{2010}^{1}$) are calculated sequentially.
4. Continue with following months and years: with the simulated 2010 consumption $D_{2010}^{1}$, storage level $s_{2010}^{1}$, and production $Pr_{2010}^{1}$, the total availability of February 2010 is known as $s_{2010}^{1} + Pr_{2010}^{1}$. Similar to January, the simulated price of February $p_{2010}^{2}$ is calculated with actual weather and consumption shock from February 2010. Repeat the procedure until December 2016, and a series of simulated data are generated with historical weather and shocks.

The historical trend of price observations is shown with the solid line in Figure 12. The historical price behaves as a downward trend curve with large fluctuation since 2010. The natural gas production in recent years has been increasing rapidly resulted from lower production cost due to the technology development. As a result, the natural gas price generally performs a downward trend as the total availability in the market increases. Besides this downward trend, price varies a lot among months, ranging from \$1.50 to \$5.80. The price peaks occur in 2010 and 2014, while the price valleys are in 2012 and 2015. With a close examination of the historical weather data, we can see that the price trend is highly correlated with the weather conditions. In

2010 and 2014, the US experienced very cold winters and the HDDs of these years are among the highest in the history. On the contrary, in 2012 the US experienced the warmest winter in the recent decade.

The monthly price mean shows a clear seasonal pattern that the price has two peaks through the year, a bigger peak in winter (January) and a smaller one in summer (July). This pattern is a good example that describes the correlation between price and weather. The price peak in winter is due to the high consumption resulting from the heating need of the cold weather and the summer peak is the result of the high cooling needs for the hot weather. The price standard deviation has a maximum value in February due to the higher consumption and lower storage level, and on the contrary, in October, natural gas price has lower standard deviation since the storage level in October is high.

Three different simulations are carried out in sequence as to investigate the effects of the weather and random shock: (1) assume weather remains stable among each year, simulation with mean HDD/CDD and no consumption shock; (2) simulation with actual historical HDD/CDD and no consumption shock; (3) simulation with actual HDD/CDD and actual consumption shock. The consumption shock is calculated as the ratio of historical consumption to the predicted consumption level using the estimated consumption function in chapter 3.

We use these three back testing simulations to see how well the competitive storage model is able to simulate price series which correspond to historical prices. Moreover, we quantify how well the price behavior can be explained with the state variables. The simulation results are assessed from different aspects to check if the model solution and simulation can fit the historical observations well, including: (1) the goodness of fit of the simulation to the historical curve; (2) the monthly statistical properties of natural gas price (mean and standard deviation); (3) the autocorrelation of natural gas price.

In order to investigate how well the model generates price series, the coefficients of determination ($R^2$) are calculated with each simulation series to show the goodness of fit of the simulation.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_i \left(p_{sim}^i - p_{data}^i\right)^2}{\sum_i \left(p_{data}^i - \overline{p_{data}}\right)^2},$$

where $i(= 1,2,\cdots,84)$ denotes the month index during the simulation period from January 2010 to December 2016: January 2010 is the 1st month and December 2016 is the 84th month. $p_{sim}^i$ is the simulated price in the $i^{th}$ month, $p_{data}^i$ is the historical price in the $i^{th}$ month, and $\overline{p_{data}}$ is the average price of all months in the simulation period.

The simulation results of all three series are displayed in Figure 12, and the monthly statistical properties are shown in Table 14 to Table 16.

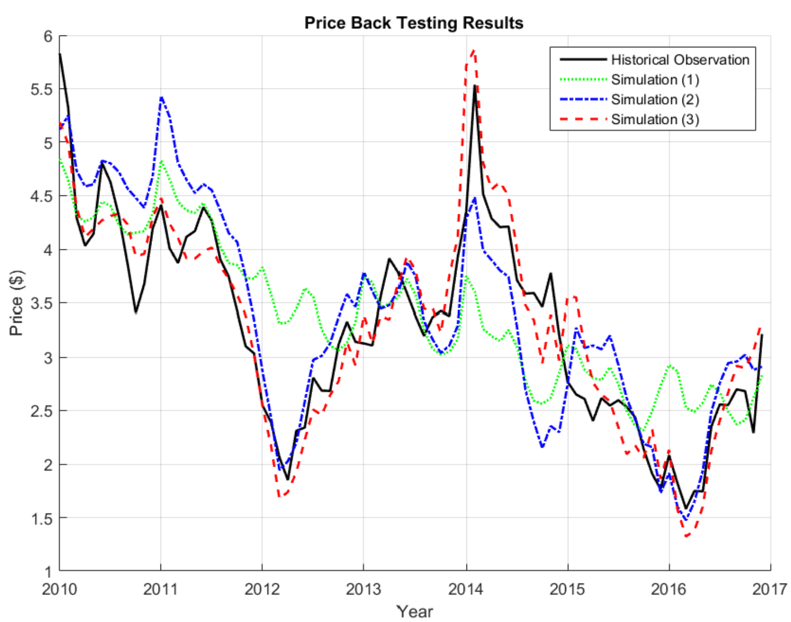


Figure 12. Price historical observation and back testing results

Note: Three simulation series display in Figure 12: simulation (1) with mean HDD/CDD and no consumption shock; simulation (2) with actual HDD/CDD and no consumption shock; simulation (3) with actual HDD/CDD and actual consumption shock.

Table 14. Monthly Price Mean Results

| Month | Historical Observation | Simulation Case 1 | Simulation Case 2 | Simulation Case 3 |
|---|---|---|---|---|
| Jan | 3.589 | 3.8628 | 3.7348 | 3.8589 |
| Feb | 3.5487 | 3.7307 | 3.699 | 3.6456 |
| Mar | 3.214 | 3.4572 | 3.3538 | 3.2548 |
| Apr | 3.1924 | 3.4165 | 3.3417 | 3.1157 |
| May | 3.2828 | 3.4481 | 3.3926 | 3.2268 |
| Jun | 3.4616 | 3.5911 | 3.6185 | 3.3737 |
| Jul | 3.4204 | 3.4759 | 3.5773 | 3.3393 |
| Aug | 3.2527 | 3.2179 | 3.3837 | 3.188 |
| Sep | 3.1995 | 3.073 | 3.2549 | 3.2123 |
| Oct | 3.0935 | 3.054 | 3.1872 | 3.0574 |
| Nov | 3.0664 | 3.116 | 3.1727 | 3.2873 |
| Dec | 3.2115 | 3.279 | 3.1026 | 3.2186 |

Table 15. Monthly Price Standard Deviation Results

| Month | Historical Observation | Simulation Case 1 | Simulation Case 2 | Simulation Case 3 |
|---|---|---|---|---|
| Jan | 1.3245 | 0.7492 | 1.3003 | 1.3311 |
| Feb | 1.452 | 0.6951 | 1.3883 | 1.5187 |
| Mar | 1.1385 | 0.7027 | 1.292 | 1.3351 |
| Apr | 1.1387 | 0.6948 | 1.1725 | 1.2116 |
| May | 1.0321 | 0.6858 | 1.0493 | 1.172 |
| Jun | 1.0475 | 0.6779 | 0.9171 | 1.0208 |
| Jul | 0.8223 | 0.6879 | 0.8194 | 0.8754 |
| Aug | 0.7059 | 0.6966 | 0.8315 | 0.8055 |
| Sep | 0.5824 | 0.7089 | 0.8252 | 0.6917 |
| Oct | 0.5051 | 0.7129 | 0.876 | 0.6032 |
| Nov | 0.706 | 0.6297 | 0.7907 | 0.53 |
| Dec | 0.7752 | 0.5775 | 0.9393 | 0.8181 |

Table 16. Price Autocorrelation Results

| | Historical Observation | Simulation Case 1 | Simulation Case 2 | Simulation Case 3 |
|---|---|---|---|---|
| AR(1) | 0.878 | 0.9398 | 0.9398 | 0.9149 |

Case 1 (Certainty of consumption): simulation series 1 represents the effects of total availability of natural gas only, assuming no consumption uncertainties from weather or other sources.

The green dot line in Figure 12 shows the simulation results. With no consumption uncertainties, the simulated price is downward sloping at a certain rate due to the production increase. A general seasonal pattern is maintained. The certainty case apparently cannot capture the extreme cases when the price dramatically increased or decreased, which may be due to high consumption caused by extreme weather or other consumption uncertainties. Therefore, the goodness of fit with Case 1 is limited and the coefficient of determination ($R^2$) is 0.48, which can be interpreted as that only 48% of the historical price variation can be explained by the total availability alone.

The statistical properties of the price are also off from actual values by assuming certainty of consumptions. The mean price shows a seasonal pattern corresponding to the seasonal fluctuation in consumption and storage (Table 14), while the price standard deviation is low and at the same level (~\$0.7) through the year (Table 15). When there is no consumption uncertainty, the price variation comes only from the production variation, which is similar for all months; as a result, we observe no seasonal pattern of price standard deviation under certainty of consumptions.

Case 2 (Uncertain weather conditions): in Case 2, actual historical HDD and CDD are involved in the simulation process. The results account for the effects of consumption uncertainty due to weather variation besides the total availability.

The blue dash-dot line in Figure 12 shows that the predicted price becomes more volatile with the introduction of weather variation. The simulation results generally follow the same pattern as the historical data, and are able to capture most of the curvature and extreme cases, especially the price drops in 2012 and 2016. In winter 2014, the simulated price serials also capture part of the increase in price due to high consumption and low availability. The goodness of fit is also improved with $R^2 = 0.67$. In other words, about 67% of the price variation can be explained by total availability and weather variation.

The statistical properties also get improved by including weather variation into the simulation. The mean price follows a seasonal pattern which is consistent with historical trend but is generally higher. The simulated monthly price variance gets elevated with the introduction of weather variation, and becomes closer to the historical observation. Price variance is still off from historical observation for month from August to December.

<u>Case 3 (Uncertain weather conditions and consumption shocks):</u> in Case 3, both actual historical HDD/CDD and the consumption shock are applied to the simulation. The historical consumption shocks are calculated as the remainders from the calibrated consumption functions, which represent the uncertainties of the consumption that is not related to price and weather variations. The consumption shock term follows a random normal distribution with mean equals to 1, and is independent with weather and price.

The red dashed line in Figure 12 shows the simulation results of Case 3. With the introduction of both weather variation and consumption shock, the price pattern better follows the historical trend, and is able to capture almost all the extreme conditions of the price increase (2014) and price drop (2012, 2016). The coefficient of determination increased to 0.85, which confirms a good fitting result with the historical observations.

The statistical properties with Case 3 agree with the historical observations. The price mean follows the same seasonal pattern with some ups and downs. The price variance also follows the historical trend that has a maximum variance in February and then decrease till the end of year.

## 5.3 Marginal Effect Analysis

As discussed above, the competitive storage model is verified via the accuracy test and the back testing simulations. The model is able to generate price series consistent with historical observations. We then can use the model solutions to investigate the marginal effects of state variables on price. The price response in different months to the state variables is different, and the price response to one variable may depend on the level of other variables, that is cross impact exist. In order to separate the impact of different state variables, only one state variable will vary at a time while the others are fixed at different levels to show the cross impact. In this section, the impacts are studied with one representative winter month (January) and one summer month (July) of 2016 to show the typical price response to the state variables.

### 5.3.1 Weather's impact on price level and variance

The impact of weather on price level is shown in Figure 13 and Figure 14, which represent the results in January and July respectively. In each graph, the solid line, dotted lines and dashed lines represent the price response given different total availability levels and consumption shocks. The solid line is the benchmark with medium total supply and no consumption shock; the dashed lines are results given no consumption shock and different total supply level; and the dotted lines represent results given medium total supply and different consumption shock level. The medium total supply levels are set to be the actual total supply values in January and July respectively (5.6 tcf in January, and 4.8 tcf in July). The lower and higher supply levels are defined as 90% and 110% of the medium supply storage of the corresponding month (5.0 tcf as lower supply and 6.1 tcf as high supply in January, 4.3 tcf as low supply and 5.3 tcf as high total supply in July). The low and high consumption shock levels are set to be 0.9 and 1.1 respectively for both January and July.

Figure 13 shows the results of price level change corresponding to heating degree days in January. Only HDD is included as CDD is close to 0 and has no impact on natural gas consumption in January. In all supply and consumption shock conditions, the price level increases along with the HDD. This is consistent with expectation: the higher HDD leads to higher residential, commercial electrical and industrial consumptions for heating purpose and sequentially results in higher price level. In addition, a lower supply level or higher consumption shock will induce higher price level as expected. The correlation between price level and HDD for all cases is close to linear with a little convex shape. The slope of the function is increasing as the supply level decreases or consumption shock increase. When natural gas supply is tight, there is pressure to reduce consumption and accumulate natural gas so as to gradually revert inventory to normal level. As a result, the price will increase more when HDD increases by the same amount under tight supply. Similarly, when consumption shock is higher, it requires more natural gas consumption when HDD increase by the same amount. Therefore the price increases more relative to benchmark case.

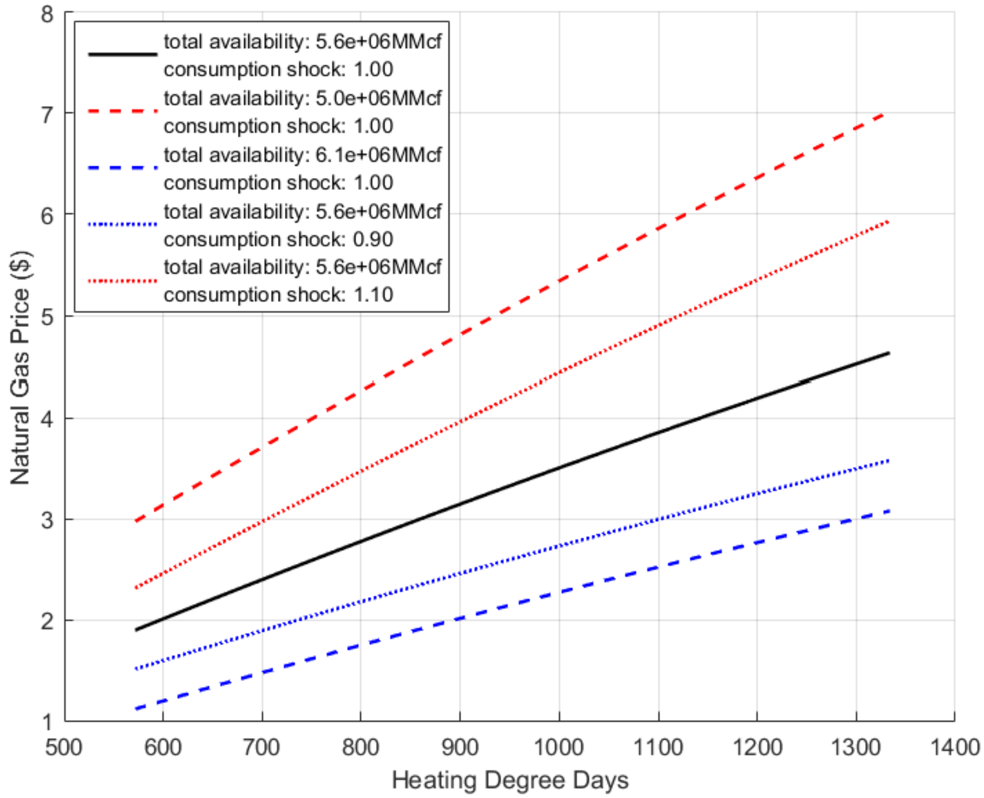


Figure 13. Weather impact on price level in January

Figure 14 shows the price sensitivity to cooling degree days in July (HDD is near 0 in July). Similar to the results in January, natural gas price increases with CDD and follows a pattern that is close to linear and with a little convex shape. The results are as expected that the price increases with the higher consumption induced by cooling purpose, mostly in electrical power sector. The slope of price function of CDD also increases with lower supply level and higher consumption shock. However, the slope change is very small and almost negligible. This is due to the fact that the production level is higher than the consumption level in July. Natural gas is plenty. As a result, the price sensitivity is lower and similar under relative tight supply level.

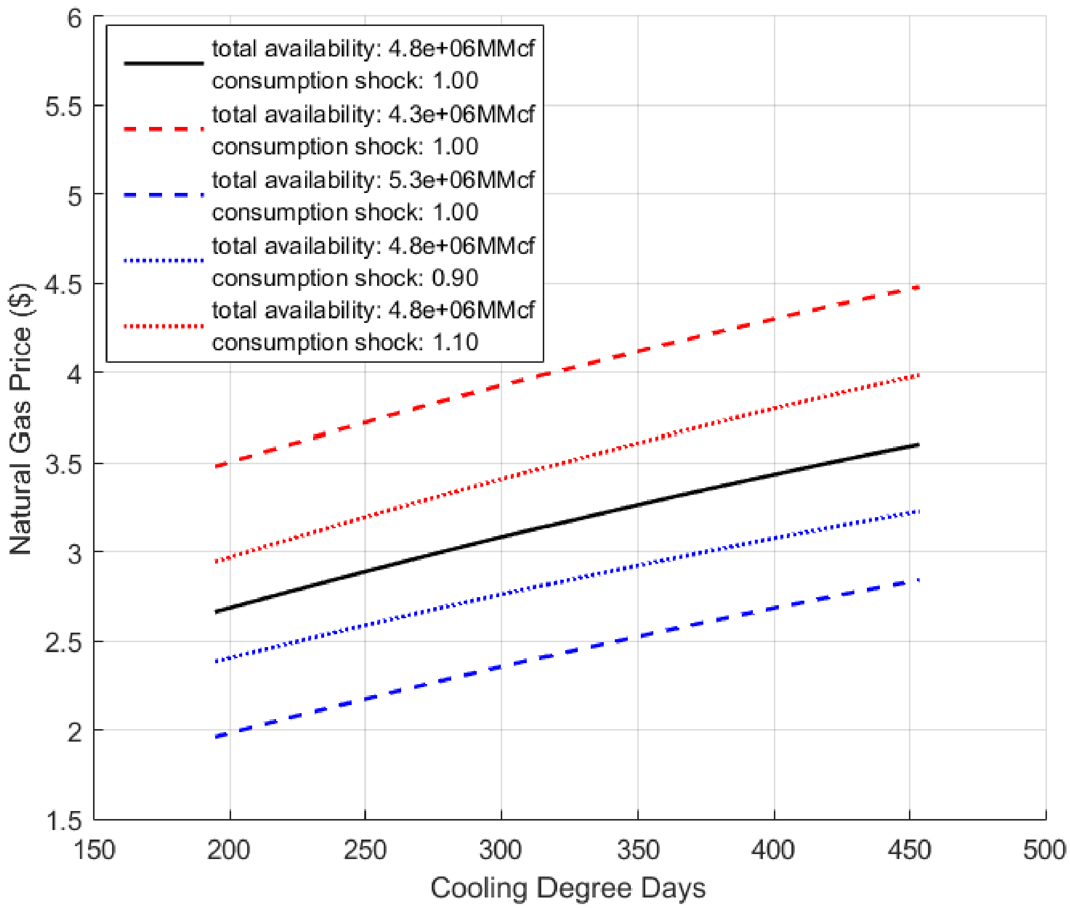


Figure 14. Weather impact on price level in July

By comparing Figure 13 and Figure 14, we can see the difference of the price response to HDD and CDD. Apparently, with the same percentage change of degree days, the impact of HDD in January is greater than that of CDD in July. This is caused by several reasons. The consumption level in January is much higher than that in July, and the price is more sensitive due to the higher needs of natural gas. Moreover, the impact of HDD on consumption is greater than CDD, since HDD will impact on all of the four sectors of consumption (residential, commercial, industrial, and electrical), while CDD mainly drives the electrical power consumption sector.

Figure 15 and Figure 16 show the weather impact on the price standard deviations of January and July respectively. By keeping the weather condition at a certain HDD/CDD level, 1000 randomly selected combinations of consumption shocks and total availability levels are used for price simulation. The standard deviation of these 1000 price level is calculated for the specific HDD/CDD level. This process is repeated with a series of HDD/CDD data, and the results are represented as the curve shown in Figure 15 and Figure 16.

The price variance increases with the weather conditions in both January and July. This is consistent with our price level observation that the price becomes more sensitive when the weather condition become extreme, especially in January. The curvature of the price standard deviation is close to linear, which is also consistent will the conclusion of the price level. In July, though the price standard deviation goes upward, the magnitude change is very small (~0.03). This is because when the supply is sufficient and consumption is low, the price volatility will not change much with consumption level. This can also be confirmed from the price level results in Figure 14 that the price function curves are close to parallel under different conditions.

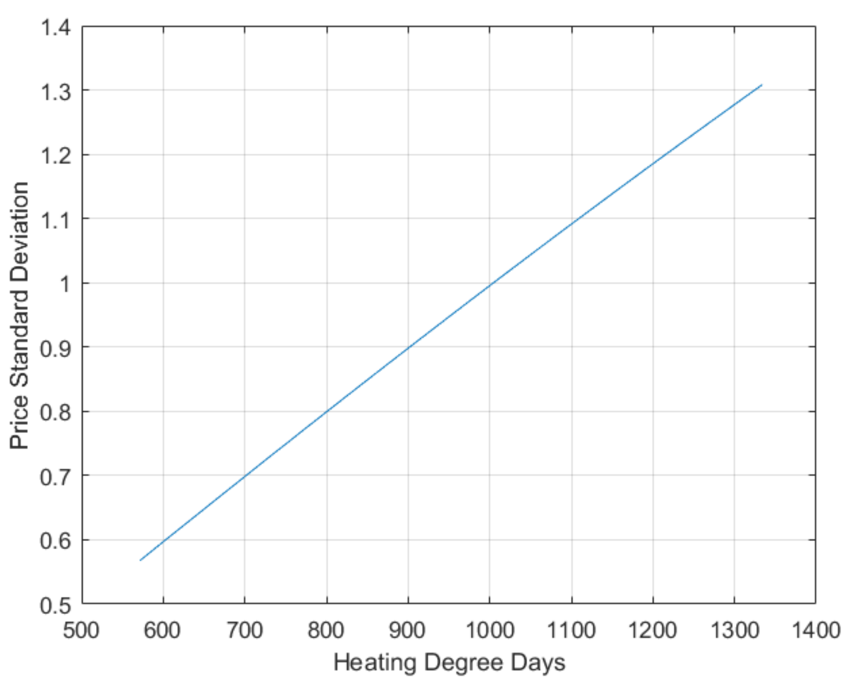


Figure 15. Weather impact on price standard deviation in January

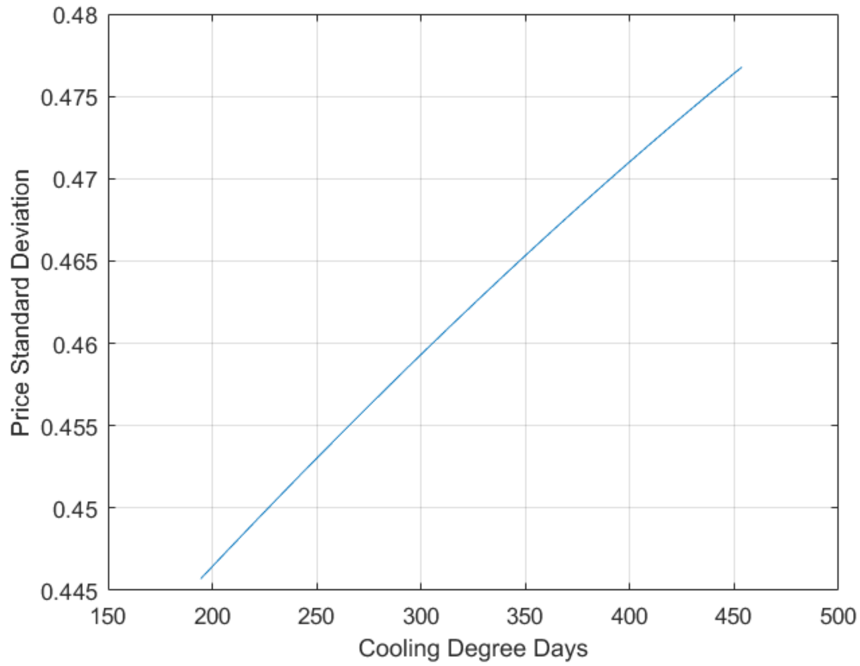


Figure 16. Weather impact on price standard deviation in July

5.3.2 Total availability's impact on price level and variance

The impact of total availability on price level is shown in Figure 17 and Figure 18 for January and July respectively. In each graph, the solid line, dotted lines and dashed lines represent the price response given different levels of heating or cooling degree days and consumption shocks. The solid line is the benchmark with medium degree days and no consumption shock. The dashed lines are results given no consumption shock and different degree days. The dotted lines represent results given medium degree days and different consumption shock level. The medium degree days are set to be the actual medium HDD in January and CDD in July (953 heating degree days in January, and 324 cooling degree days in July). The lower and higher degree days are defined as 90% and 110% of the medium level of degree days of the corresponding month, that is 857 and 1048 HDD in January, 292 and 356

CDD in July. The lower and higher consumption shock levels are set to be 0.9 and 1.1 respectively for both January and July.

Figure 17 shows the price as a function of natural gas total supply under different weather and consumption shock conditions in January. The price level decreases with total supply. Under same supply level, higher price level is observed with higher HDD or higher consumption shock. The price change is material with a +/-10% change of the total supply when HDD or consumption shock is high. The price functions follow a convex curvature such that the slope of the price functions keeps falling when total supply increases. When total supply reaches a certain high level, natural gas price becomes less sensitive to the supply changes in the market. The marginal effect decreases. When total availability is low, the reduction in natural gas supply will introduce higher price change. A higher HDD or higher consumption shock will lead the price function curvature to be more convex. Both a higher HDD and a higher consumption can lead to higher total consumption. With higher consumption, the lack of total supply will lead to a more "panic" market and make the price more volatile.

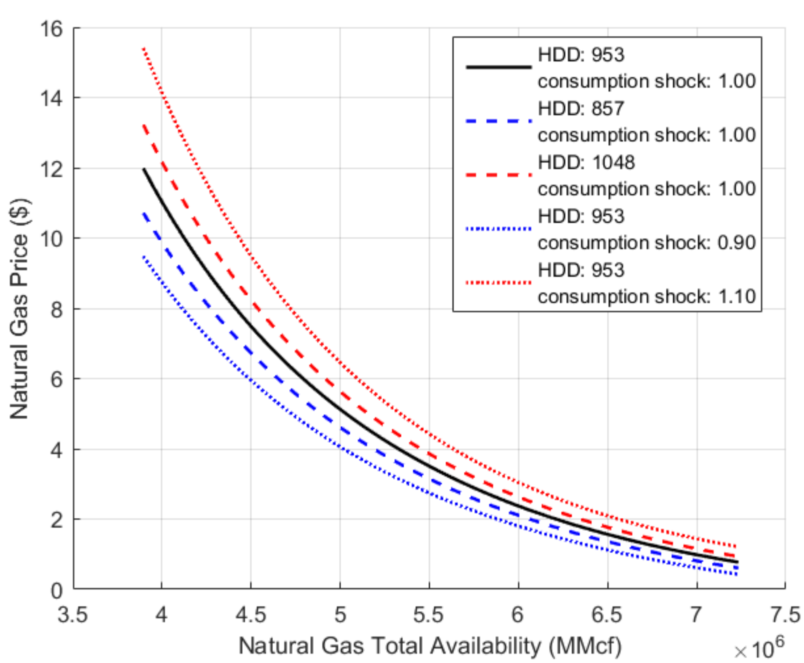

Figure 17. Total availability impact on price level in January

Figure 18 shows the price response to natural gas total supply in July. Similar to what we observed in January, the price functions follow a downward trend pattern with a convex curvature. Higher price level is observed with higher CDD or higher consumption shock. As total supply changes, price ranges from $1 to around $6. The price sensitivity reduces when the total supply increases due to the same reason as in January. The curvature of the functions does not change much when introducing CDD or consumption shock. The curve simply moves up and down with different weather and shock conditions. In July, production level is higher than the total consumption and this makes the production sufficient. The consumption change induced by random shock or different weather conditions is not large enough to change the price sensitivity. As a result, at the same total supply level, the shift of CDD or consumption shock will only affect the price level but not much on the price sensitivity.

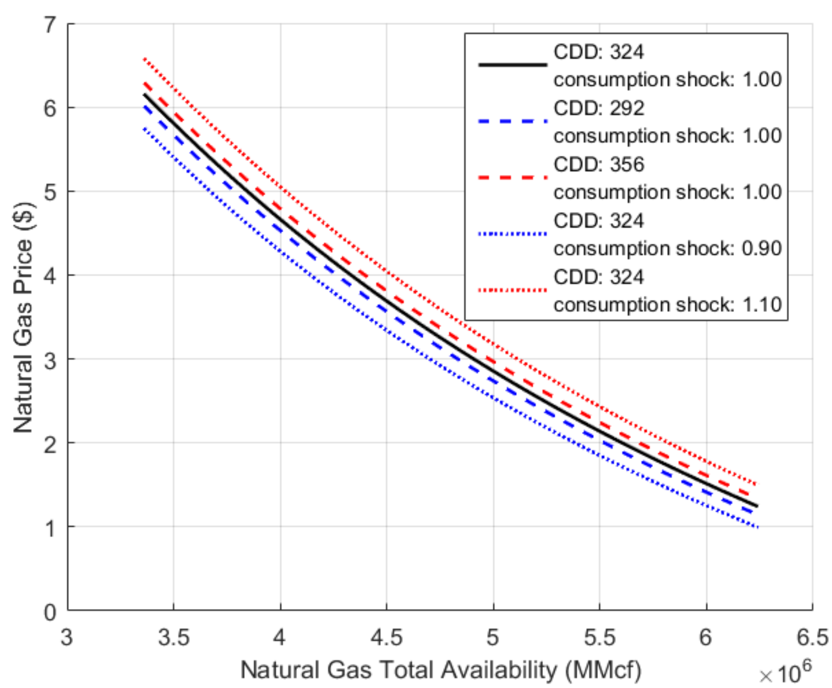

Figure 18. Total availability impact on price level in July

By comparing the price functions in January and in July, we can observe that the price response to natural gas total supply is much greater in winter than that in summer. The higher consumption in winter makes the supply-demand relationship tight and market is more volatile. Price is more sensitive especially at a lower supply level. Extreme weather conditions and high demand shock have material impact on the price sensitivity in winter, while they have less impact in summer.

Figure 19 and Figure 20 show the total availability impact on the price standard deviations of January and July respectively. Simulation process similar to weather impact investigation is applied. 1000 combinations of consumption shocks and weather conditions are randomly selected and the price standard deviation is calculated afterwards.

The price standard deviation drops as the total supply increases in both January and July, since higher total supply leads to a more stable market. The magnitude of price standard deviation in January is higher than July due to the high consumption level. Moreover, the curvatures of the two functions are quite different that price standard deviation function shows obvious convex pattern in January, while the curve is close to linear in July. In January, the total availability has significant impact on the price variance. When there is lack of total supply, natural gas price varies a lot if there is weather or consumption shock. In July the total availability impact is weaker than in January. Lower supply level will lead to a higher price standard deviation, but the change is limited due to the low consumption level.

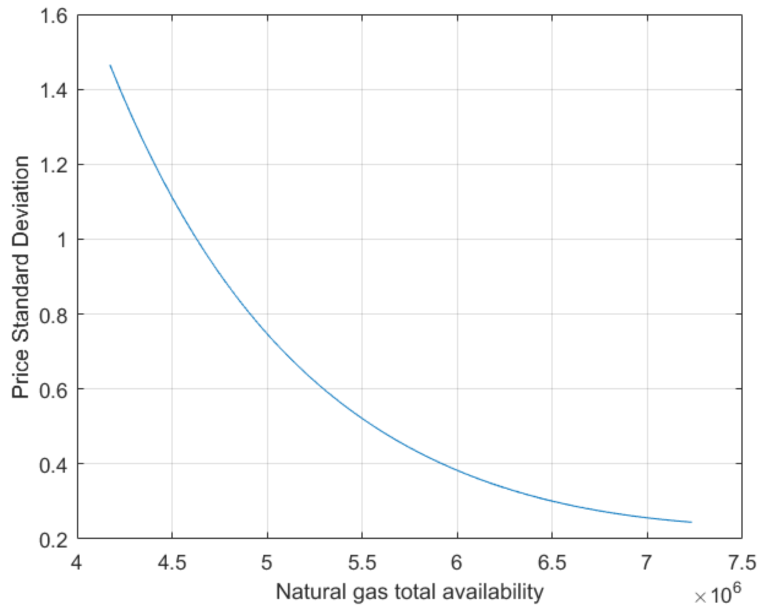


Figure 19. Total availability impact on price standard deviation in January

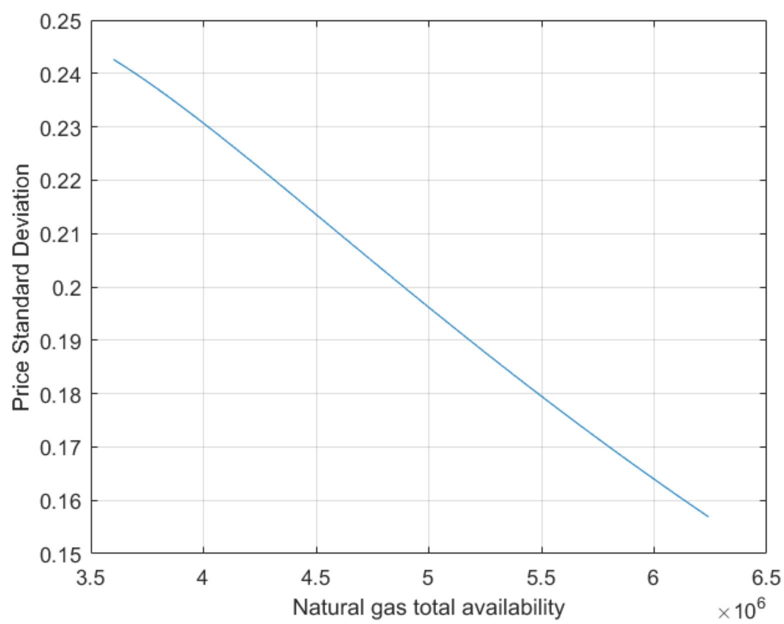


Figure 20. Total availability impact on price standard deviation in July

# CHAPTER 6. US DOMESTIC PRICE AND VOLATILITY RESPONSE TO LNG EXPORT

This chapter employs the model constructed in previous sections to assess the impact of potential LNG export on U.S. domestic market.

## 6.1 Introduction

The rapid ramp-up of domestic natural gas production reshaped the U.S. energy sector for the most recent decade. Natural gas production increased significantly because of the shale gas boom facilitated by the technological advances of hydraulic fracking and horizontal drilling starting in 2005. In 2005 U.S. natural gas supply was 18,927 billion cubic feet with negligible gas withdrawal from shale gas wells. Supply reached 28,752 billion cubic feet in 2015 with 47% of the production coming from shale gas wells. Total production increased by 51.91% in 8 years. The increased domestic production is mostly used for increased domestic consumption and partly displaces natural gas imports from Canada, Trinidad and Qatar. Natural gas is increasingly used for electricity generation, transportation fuel, and some industry feedstock. Two of the most prominent changes are substitution of natural gas for coal in the power generation sector and the emerging opportunity of natural gas export.

The United States is gradually transforming itself from a net natural gas importer to a potential top exporter. U.S. is a net natural gas importer with imports coming primarily from Canada and Mexico through pipelines, and some from Trinidad in the form of liquefied natural gas (LNG). U.S. also exports natural gas to Canada and Mexico through pipelines and to Asian countries such as Japan in the form of LNG. U.S. natural gas export increased dramatically in the beginning of the 2000’s, while imports decreased since 2005 when the shale gas production started to increase. As a result, net import decreased substantially since 2005 (Figure 21).

Although U.S. is still a net natural gas importer until 2016, with net import of 671 bcf in 2016, U.S. has already become a net LNG exporter of 98 bcf. Most of the LNG is exported to Japan. EIA (2017) predicts that U.S. will become a net gas exporter in 2017. By 2020, when all current U.S. liquefaction projects are expected to come online, the U.S. will comprise almost one-fifth of world liquefaction capacity and will become the third-largest LNG export capacity holder in the world, after Qatar and Australia. This thesis focuses on LNG export because currently the natural gas market is quite isolated due to high shipping cost, large investment and market power. LNG is the major format for global trade because major Asian markets cannot be reached by pipelines.

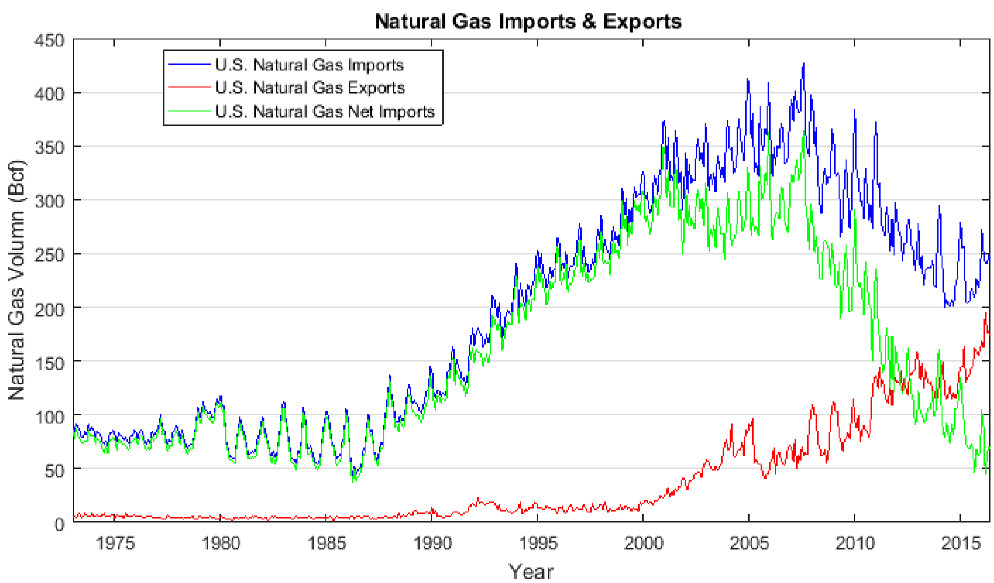


Figure 21. U.S. Natural Gas Imports, Exports from 1973 to 2016

Data Source: EIA

Large volumes of LNG export is likely to be realized in the near future due to current large pricing spread between the U.S. and the rest of world especially Asian countries, along with recent policy promotion and upcoming facility completion. The U.S. natural gas market is quite isolated due to high shipping cost, large investment and market power. Price differs significantly from region to region. As shown in Figure 22, the U.S. natural gas price has fallen dramatically because of increased domestic production since 2005 from $10 to about $3 recently. The price in Europe and Japan remains at a high level compared to the US price. European natural gas price averaged $4 higher than in the U.S. from 2005 to 2016. The price difference is even larger between US and Japan, the largest LNG importer of Asia, at $5.80 on average.

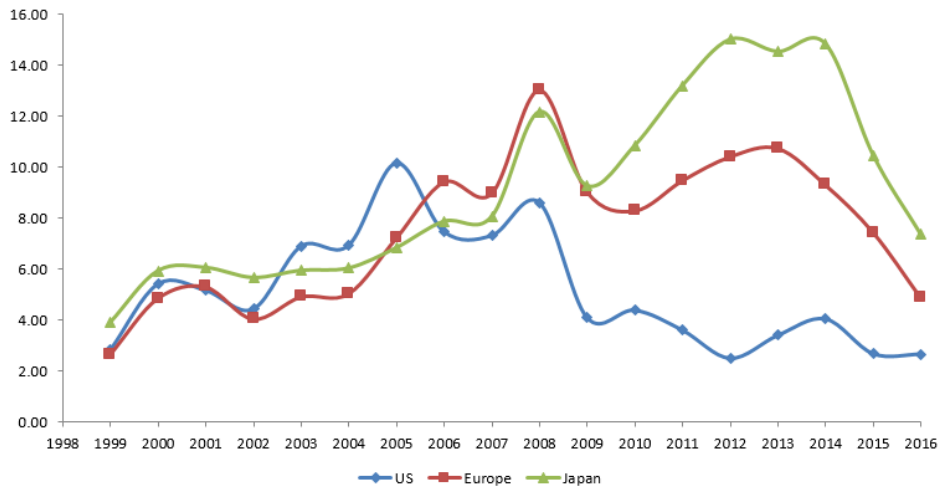


Figure 22. Natural Gas Price ($/MMBtu) Comparison between U.S. and Other Regions

Data Source: World Bank http://www.worldbank.org/en/research/commodity-markets

Beside the economic incentive, in May 2017, the Trump administration signed trade deals with China and one main item is to facilitate LNG exports to China. China is one of the fastest growing LNG markets and is potentially a large LNG importer. China has no free trade agreement with U.S. and its U.S. LNG demand is suppressed. Chinese companies currently have no long-term contracts with U.S. natural gas suppliers. The Energy Department has authorized up to 19.2 bcf per day natural gas shipments to China and other interested countries under the trade deal. In addition to the policy incentive, the recent online LNG export terminal makes gas export feasible. In early 2016, Cheniere Energy's Sabine Pass export terminal in Louisiana has been put into production. U.S. export capacity is expected to continue to expand rapidly in the following three years with several projects under construction like Cove Point, Cameron and

Freeport. The expansion of the Panama Canal in June 2017 made the shipment of LNG much easier and less expensive.

The potential for large U.S. LNG exports has triggered much debate among policy makers and market participants. From a policy perspective, it is required that any U.S. natural gas export entity should apply for permission from the U.S. Department Energy (DOE). By the Natural Gas Act, if the natural gas is exported to a country that has free trade agreement (FTA nations) with the U.S., DOE must grant the permission. If the export destination is to a non-free trade agreement country, according to Natural Gas Act, the opposing party has the burden to prove that the proposed export is not in the public interest. If the party who opposes the natural gas export fails to prove the export is against the public interest, DOE must grant the license for export. There are over 20 applications being submitted to DOE for approval.

Along with the large number of LNG export applications submitted to DOE, there are widely discussed concerns and questions regarding the impact of large-scale LNG export on the U.S. domestic market. The parties opposing export are worried that the United States natural gas price will increase greatly and hurt domestic consumers, as well as gas related industries. The impact on the environment and climate change is one of the major concern as well (Levi, 2012; Ratner et al, 2015). The proposing parties believes that natural gas export will create direct and indirect job opportunities. In addition, LNG export ensures continuous and sufficient investment in natural gas industry and promote related jobs. These benefits overshadow the negative impact from the slight price increase according to export supporters (EIA, 2014; DOE, 2015; Levi, 2012).

Theoretically, exporting LNG increases total gas demand and will increase natural gas price. The price spread will tighten between US market and the rest of the world, especially LNG importing countries like Japan and China. Increased market price benefits market participants associated with gas production. High natural gas price stimulates investments and fosters ongoing natural gas production increase. On the other side, increased natural gas price is harmful for natural gas users. Residential and commercial gas users' welfare will decrease. It might slow down the pace of power generator coal-to-gas switching. Increased price might also be detrimental for industrial users like fertilizer companies and other manufacturers.

DOE initiated two studies to understand the potential impact of allowing large-scale LNG export to non-FTA nations: the Energy Information Administration (EIA) performs one of the two studies in 2014 and external entity conducts the other one released in 2015. Both studies conclude that LNG export is overall beneficial for U.S. economy. The gain from international trade overweighs the loss incurred by higher natural gas price.

The viability and potential impact of large-scale LNG export are widely discussed in the current literature. Medlock (2012) utilizes an international trade analytical framework and argues that the impact of export will not be large because the long-term export volume is not likely to be large given the potential global natural gas market development. BP Energy Outlook 2017 predicts that the U.S. LNG export will increase dramatically and becomes one the world top three exporters. Levi (2012) proposed to analyze LNG export impact from six dimensions: macroeconomic metrics like jobs and balance of trade, distributional welfare impact, national oil security, climate change, foreign and trade policy, and local environment. Levi finds that if there is appropriate environment protection policy and the U.S. is able to leverage exports to promote broader trade, it is very likely that the benefit of export is bigger than the costs. Deloitte (2011) assumes 6 bcf/day export volume. Its integrated North American Power, Coal and World Gas Model projects that from 2016 to 2035, the weighted average price impact is $0.12 per MMBtu.

Moryadee, Gabriel and Avetisyan (2014) use the World Gas Model to investigate the potential effects. They find the average domestic price increases by about 10.9% and the natural gas price of Europe and Asia decreases significantly. Bernstein, Tuladhar and Yuan (2016) analyze LNG export under several assumptions about U.S. natural gas resource outlook, regulation environment and changing geopolitical conditions, using a global natural gas market model. The realized export level is subject to different assumptions. They believe the market is self-correcting and propose to leave it to the market to determine the volume and destination of LNG export.

As discussed above, the impact of future LNG exports on the domestic natural gas price has been the topic for many economic research studies. However, there is little discussion or analysis about how LNG exports will impact natural gas price volatility and whether the current seasonal price pattern will change because of LNG exports. When there is large-scale natural gas export, the United States natural gas market will link more closely with the rest of the world. Will the market be more volatile or less volatile than now? Furthermore, how domestic natural gas storage will respond to the new market structure remains unanswered.

This study aims to deepen the analysis in the current literature and get a better understanding about the domestic impact of US LNG export, using the monthly rational-expectations storage model developed in previous chapters. The monthly rational-expectations storage model is useful for analyzing the LNG export impact mainly due to two reasons. First, the inclusion of storage makes the model dynamic and serially correlated. The price in each period is inter-temporally linked by the non-arbitrage condition and inventory. This feature enables us to see how domestic natural gas price and volatility evolves over time. Assuming LNG export, the instantaneous and long-term response of the market are different. In the short-term, if there is large volume of gas export, price will increase significantly. But in the long run production will expand and domestic gas consuming sectors will take actions to reduce the negative impact incurred by large LNG export and utilize other alternative energy source or adopt energy efficient appliance and equipment. As all those actions take time to implement, the magnitude of the price increase will reduce gradually. Secondly, as foreign countries' gas consumption patterns throughout one year is different from the U.S., the impact of gas export for each month or different season might be different. The utilization of a monthly model is able to assess the impact on each single month.

The rest of this chapter is organized as follows. The different kinds of potential LNG export patterns or contract type are explained in section 2. Section 2 also discusses briefly how to apply the competitive storage model to assess LNG export impact and the solving and simulation algorithm. Section 3 analyzes the model result. Section 4 discusses the uncertainties faced by the U.S. export suppliers and concludes.

## 6.2 Model and Solution Algorithm

This section describes different possible LNG export contract features and how the export sector can be modelled to be part of the rational-expectations competitive storage model. The solution and simulation algorithm are explained in this section as well.

### 6.2.1 Contract Type: Fixed Volume or Endogenous Volume

Most future U.S. natural gas exports will take the form of LNG as no pipeline is available between U.S. and other major importing countries. LNG can be shipped to different parts of the world in response to different regional production and demand relationship. U.S. LNG export is quite promising given the large pricing spread shown in Figure 22 and the fast growing natural gas demand around the world, especially in Asian nations. The top two importing regions are Europe and Asia. Europe has domestic natural gas resources and is much less dependent on import compared to Asia. Given the pricing spread and growth potential, Asia is most likely to become the major importer of U.S. LNG. The top natural gas importers around the world are Japan, South Korea and China, accounting for over 50% of the global LNG trade in 2015, with total average of 18.2 bcf per day[6.] China is expected to have increasing natural gas demand due to more rigorous environment policies and growing LNG importing capacity. The natural gas demand of Japan and Korean is expected to increase as well due to the recent shift from nuclear to gas-fired power generations. China natural gas import increased from 0.1 tcf in 2007 to 2.2 tcf in 2014. Japan's import volume increased from 3.5 tcf in 2007 to 4.7 tcf in 2014[7]. The increasing trend is expected to persist.

The U.S. market participants have started to take advantage of the attractive pricing spread and potential profit already. As of May 2017, there are 11 LNG export terminal projects approved by Federal Energy Regulatory Commission (FERC). The total export capacity is 16.44 bcf per day. Out of the 11 terminals, 7 projects are under construction with total export capacity of 9.56 bcf per day[8].

Before assessing the impact of LNG export, we need to figure out what are the possible export scenarios and how LNG export contracts are going to be specified. As shown in Figure 22, the United States natural gas price is divergent from Europe and Asian countries. Unlike global oil market, which has a global benchmark price, worldwide natural gas markets are isolated. The natural gas price in the U.S. is market driven, based on the total supply and

---

[6] Data source: EIA (2016), https://www.eia.gov/todayinenergy/detail.php?id=27652
[7] Data source: U.S. Energy Information Administration, International Energy Statistics
[8] Data Source: FERC https://www.ferc.gov/industries/gas/indus-act/lng/lng-approved.pdf

consumption level. In Europe and Asia, the natural gas price is mainly indexed to crude oil prices and will not necessarily reflect the actual supply and demand relationship. The two different pricing mechanisms make the regional natural gas prices disconnect from each other. Figure 23 displays the relationship between Brent Crude Oil Price and the natural gas price in Europe and Japan. They are highly correlated with each other. The correlation coefficient of Brent Crude Oil price and Europe price is 0.86 using the monthly data starting from 2000. The correlation coefficient of Brent Crude Oil price and Japan natural gas price is 0.87. The correlation coefficient between Brent Crude Oil and U.S. natural gas price is quite small, just 0.1 using the monthly data starting from 2000 to 2017.

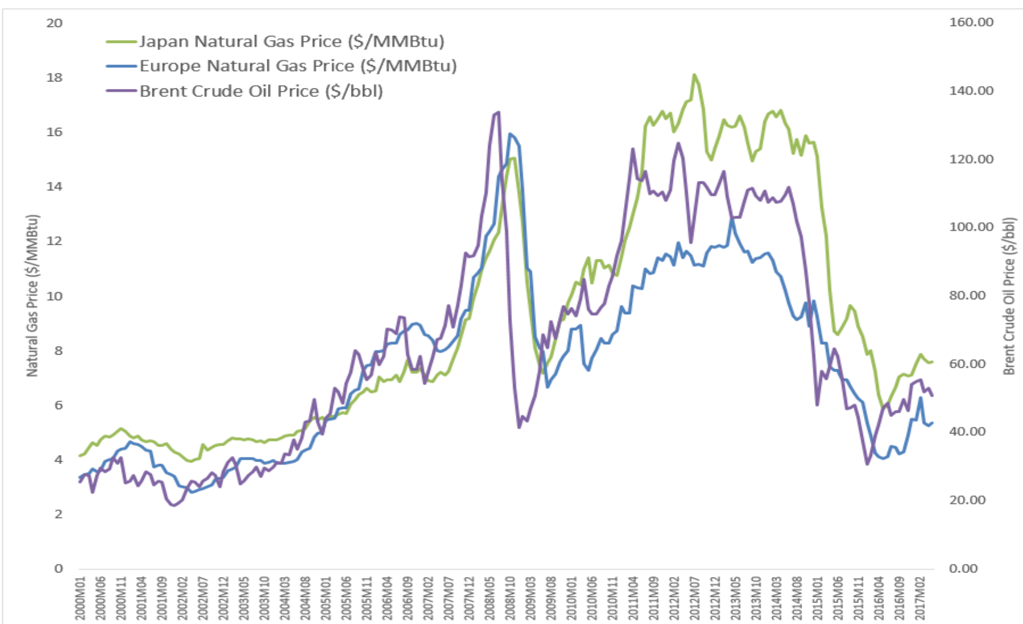


Figure 23. Correlation between Natural Gas Price and Brent Crude Oil Price

Source: World Bank http://www.worldbank.org/en/research/commodity-markets

The difference in pricing regime among regions mainly resulted from limited LNG liquefaction and re-gasification capacity, high transportation cost, limited transportation capacity along with policy regulations. To export natural gas globally, natural gas must be cooled and converted from a gaseous to liquid state to decrease the volume to about 1/600$^{th}$ of the original volume The shipment of LNG requires special refrigerated storage tanks and special LNG vessels. When LNG reaches its export destination, it is returned to its original gaseous state. The whole liquefaction – shipping – regasification process is expensive, with commonly estimated cost of ~$6 per MMBtu to ship natural gas from the United States to Asian countries (Baron, 2014). The LNG related facilities are capital intensive. It takes around $5 to $10 billion for each LNG export project.

The above features of global natural gas market shape the LNG contract. Most of the LNG contract are long-term contracts. Meanwhile, the global natural gas is gradually being transformed from being isolated to being integrated in recent years. This study designs two broad types of LNG export contract. One type of contract has importer purchasing a fixed predefined amount of natural gas from the United States, regardless of the natural gas price level. The other type of contact does not specify the volumes exported. The LNG export amount is dependent on the transportation cost and pricing spread between the United States and the importing county.

The fixed export amount scenario is meant to be consistent with EIA's 2014 report and the follow-up studies based on this report. EIA performed an assessment of LNG export impact on the domestic market in response to the request from the U.S. Department of Energy's office of Fossil Energy (DOE/FE) in 2014. DOE/FE specified three export scenarios in its request: 12 Bcf per day, 16 Bcf per day and 20 Bcf per day, with a phase-in rate at 2 Bcf each year. EIA believes that the export ramp up speed and the ultimate 12 Bcf per day scenario is extremely aggressive and almost impossible to realize. They are intended to show an outer envelope result following the approval of export licenses. Deloitte performed an impact analysis using its integrated North American Power, Coal, and World Gas Model in 2012. The results are based on 6 Bcf per day LNG export.

Some of the potential U.S. LNG export contracts will assume fixed export amount regardless of price level if the LNG is used for baseload power generation or for energy security reason. Asian countries have limited natural gas resource and few LNG suppliers. They are eager to diversify their import source to secure gas supplies and gain more negotiation power via the source diversification. It is observed from current market that the major importers, especially Asian countries place high priority on security of gas supply (Ritz, 2014; Vivoda, 2014). They tend to sign long-term contract with predefined volumes to avoid supply shortage. Those importers will buy US LNG export even though it is not economically competitive compared to other suppliers. In addition, as shown by Alterman (2012) and Pindyck (2004), natural gas price is much more volatile than oil price. Securing long-term supply can reduce cash flow variability and is beneficial for financing.

Although long-term natural gas contracts indexed to oil price dominate the current market, both Asian and European nations are gradually moving from oil-indexed contract to gas hub based pricing. It is foreseeable that global natural gas market is going to be more competitive and integrated along with development and expansion of international LNG trade. International Energy Agency World Energy Outlook 2011 predicts that the percentage of LNG indexed to Crude will decrease 82% to 63% in 2035. Therefore, we also utilize the international trade framework and analyze LNG export scenarios where price is endogenously determined by competitive markets.

Cheniere Energy, the leading LNG export company has signed several LNG contracts to export U.S. domestic natural gas to Japan, Korean and India. The price of these contracts is indexed to Henry Hub natural gas price, with a 15% premium and a fixed fee to cover operating costs. The contracts do not have terms regarding binding commitment to take gas. However, the buyers are required to pay the fixed fee whether they take the gas or not. According to EIA, about 80%[9] of the LNG export volume for the projects under construction is directly indexed to Henry Hub price or indirectly indexed to Henry Hub via a hybrid pricing mechanism. The widely adopted gas hub pricing mechanism will promote the transition to trading hubs pricing globally.

### 6.2.2 Model Framework and Solution Algorithm

The model setup for endogenous LNG export is similar to Makki (1996).

U.S. market clearing condition

The previous discussed competitive storage model needs to be modified slightly to include the existence of LNG exports. In the U.S. market, the total availability of natural gas is not only consumed by domestic customers as the residential, commercial, industrial, and electricity consumptions discussed before, but also as a source for export purpose. The domestic market clearing condition is hence changed so that export quantity is added as part of the total demand.

$$TP^{(m-1)} + s^{(m-1)} = TS^{(m)} = D_{tot}^{(m)} + X^{(m)} + s^{(m)},$$

In the equation above, $X^{(m)}$ denotes the total natural gas export. The other terms still have the same definition. Total supply ($TS^{(m)}$) comes from the production and inventory carried over from last month. The total consumption ($D_{tot}^{(m)}$) is the summation of the four consumption sectors as functions of the U.S. natural gas price. The inventory level ($s^{(m)}$) is the leftover which is used to balance the monthly production-demand gap.

[9] Source: EIA (2015) https://www.eia.gov/todayinenergy/detail.php?id=23132

When natural gas export is introduced to the market, the demand of natural gas will increase. The increased demand will lead to an increase of the U.S. natural gas price and sequentially lead to a higher production and lower domestic consumption. The market clearing condition will be achieved at a new equilibrium state.

Export determination

In this study, both exogenous and endogenous scenarios of export are investigated. In the exogenous scenario, the daily export is simply a fixed amount and does not change with the US domestic natural gas price. The monthly export quantity is then fixed as

$$X^{(m)} = x \cdot days^{(m)},$$

where $x$ is the fixed daily export amount and $days^{(m)}$ is the number of calendar days in month m.

Under endogenous LNG export scenarios, the export quantity is determined by the US domestic natural gas price and the price of importing countries.

$$\begin{cases} p^{(m)} + \tau = p_{im}^{(m)}, & if\ X^{(m)} > 0, \\ p^{(m)} + \tau > p_{im}^{(m)}, & then\ X^{(m)} = 0, \end{cases}$$

In the equation above, $p^{(m)}$ is the U.S. domestic natural gas price. $p_{im}^{(m)}$ is the price of the importing country and $\tau$ represents the full shipping and handling cost including all costs or benefits incurred by LNG trade besides the natural gas price, such as shipping, gas liquefaction and government export subsidy.

The simple international trade condition describes the competition between the US and other exporting countries. From importing countries' side, if the total export price from US (US domestic price plus shipping) is higher than the market price of the importing countries, the US price will not be competitive and no trade will take place. Otherwise, export will happen and the price of US and the importing countries will be related by above equation.

It is assumed that the importing country has no inventory, which reflects the reality for most Asian importers. As a result, the market clearing condition of the importing countries is that the monthly total consumption ($D_{im}^{(m)}$) equals to the total imports from U.S. ($X^{(m)}$) and import from other countries or domestic production ($X_{other}^{(m)}$).

$$D_{im}^{(m)} = X^{(m)} + X_{other}^{(m)}.$$

The total consumption of the importing countries and the export from other countries are defined as functions of the market price of the importing countries ($p_{im}^{(m)}$) with estimated elasticities. The consumption ($D_{im}^{(m)}$) is a downward-sloping function of price and the export/domestic production ($X_{other}^{(m)}$) is an upward-sloping function of price:

$$D_{im}^{(m)} = D_{im}^{(m)}\left(p_{im}^{(m)}\right) = \alpha_D^{(m)}\left(p_{im}^{(m)}\right)^{\beta_D},$$

$$X_{other}^{(m)} = X_{other}^{(m)}\left(p_{im}^{(m)}\right) = \alpha_S^{(m)}\left(p_{im}^{(m)}\right)^{\beta_S},$$

where $\beta_D < 0$ is the demand elasticity of importing countries, and $\beta_S > 0$ is the supply elasticity of other exporting countries. Based on the international trade condition, the LNG export from the U.S. can be expressed as

$$\begin{cases} X^{(m)} = \alpha_D^{(m)}\left(p_{im}^{(m)}\right)^{\beta_D} - \alpha_S^{(m)}\left(p_{im}^{(m)}\right)^{\beta_S} = \alpha_D^{(m)}\left(p^{(m)} + \tau\right)^{\beta_D} - \alpha_S^{(m)}\left(p^{(m)} + \tau\right)^{\beta_S}, & if\ X^{(m)} > 0; \\ X^{(m)} = 0, \quad if\ p^{(m)} + \tau > p_{im}^{(m)}. & \end{cases}$$

Or equivalently,

$$X^{(m)} = max\left\{\alpha_D^{(m)}\left(p^{(m)} + \tau\right)^{\beta_D} - \alpha_S^{(m)}\left(p^{(m)} + \tau\right)^{\beta_S}, 0\right\}.$$

Modified competitive storage model and solution algorithm

With the modified marketing clearing condition and the export functions for exogenous export and endogenous export, the model is closed. The new equation set is defined as below with the inclusion of the export function.

$$\frac{1}{1+r}E_m\big(p^{(m+1)})\big) = p^{(m)} + sc'(s^{(m)}) \tag{27a}$$

$$E_m\big(p^{(m+1)})\big) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{5} w_{jk}\hat{p}^{m+1}\left(\begin{matrix} hdd_i^{(m+1)}, cdd_i^{(m+1)}, \varepsilon_{D,j}^{(m+1)}, \\ Pr\Big(Pr^{(m-1)}, E_m\big(p^{(m+1)}\big), P_{mean}^{(t)}, cost^{(t)}\Big)\varepsilon_{Pr,k}^{(m)} + s^{(m)} \end{matrix}\right) \tag{27b}$$

$$TS^{(m)} = \sum_{i=1}^{4}\sum_{j=1}^{9} D_{ij}^{m}\Big(D^{(m-1)}, p^{(m)}, hdd_j^{(m)}, cdd_j^{(m)}\Big)\,\varepsilon_D^{(m)} + X^{(m)} + s^{(m)} \tag{27c}$$

$$X^{(m)} = \begin{cases} x \cdot days^{(m)}, & exogenous\ export \\ or, max\left\{\alpha_D^{(m)}\big(p^{(m)}+\tau\big)^{\beta_D} - \alpha_S^{(m)}\big(p^{(m)}+\tau\big)^{\beta_S}, 0\right\}, & endogenous\ export \end{cases}. \tag{27d}$$

With the equation set above, the price function is solved by the numerical iteration method with Chebyshev polynomial approximation. The solving algorithm is the same as discussed in Chapter 4.3:

1. Chebyshev polynomials and nodes of each state variable are defined with each month and each region. The price function is then approximated as equation (14) with the Chebyshev polynomial and corresponding coefficients.

2. The initial guess of coefficients ($c_0^{(m)}$) of the 12 months are made, with the initial guess the equation sets are solved at each of the state variable nodes and the solution $p_0^{(m)}$is calculated at all iteration nodes for each month.

3. The coefficients are updated to $c_1$ with the calculated $p_0$ as $c_1^{(m)} = \left(\Phi_{N\times N}^{(m)}\right)^{-1} p_0^{(m)}$.

4. If the difference between $c_1^{(m)}$ and $c_0^{(m)}$ of all months is smaller than the predetermined convergence level, then $c_1^{(m)}$ is the solution and price approximation is solved. Otherwise repeat steps 2 and 3 with the new coefficient guess $c_1^{(m)}$ until the coefficients coverge.

### 6.2.3 Scenarios and Model Parameterization

To solve the above competitive storage model with LNG export component, we need to specify and calibrate the parameters first. This section describes in detail the scenarios this study analyzes. Model parameters are calibrated and discussed as well. We use Japan as the representative importing country in this study.

As discussed above, two broad categories of LNG export scenarios are analyzed: exogenous export with fixed volumes and endogenous export volumes. In addition to the scenarios analyzed, this study analyzes the sensitivity of results to some key parameters, including the demand and supply elasticity of importing countries and the U.S. domestic supply elasticity. All the LNG export scenarios are listed in Table 17.

For the fixed export scenarios, we need to specify the export volumes. To be comparable with EIA's report and other existing literature, there are two export scenarios: 6 bcf per day and 12 bcf per day. Both scenarios have a phase-in speed of 2 bcf each year, starting from 2017.

For endogenous export scenarios, the following functions or parameters need to be calibrated and specified: 1) consumption function of the importing countries; 2) supply function of the importing countries. The supply includes the import from other countries and domestic production, excluding those imported from the U.S.; 3) shipping cost.

Table 17. U.S. LNG Export Scenarios

| Category | Scenarios |
|---|---|
| | Scenario 1) 6 bcf/day, phase in at 2 bcf each year<br>Scenario 2) 12 bcf/day, phase in at 2 bcf each year |
| | Scenario 3) no growth, current Asia demand and supply structure remains<br>Scenario 4) Asia LNG demand doubled by 2035 and LNG supply other than US increased by 1.5 times, linear growth |
| Asian LNG Supply Elasticity | Scenario 4.1) Asian LNG supply elasticity decreases by 50% compared to Scenario 4<br>Scenario 4.2) Asian LNG supply elasticity increases by 100% compared to Scenario 4 |
| Asian LNG Demand Elasticity | Scenario 4.3) Asian LNG Demand Elasticity increases by 100% compared to Scenario 4<br>Scenario 4.4) Asian LNG Demand Elasticity increases by 200% compared to Scenario 4 |
| U.S. Domestic Supply Elasticity | Scenario 4.5) $\beta_1$ decreased by 50%<br>Scenario 4.6) $\beta_1$ doubled |

Due to data availability, this study uses Japan, the largest LNG importer, as the representative importing country. Its consumption and supply is calibrated and scaled to represent the entire Asian area. The monthly consumption data is from Japan Ministry of Finance. Constant elasticity demand and supply function are assumed for Japan. In the base scenarios, the demand elasticity is set at -0.06 and the supply elasticity is set at 0.1, which is consistent with the NERA Economic Consulting (2014). Asian countries usually have lower demand elasticity than U.S. due to energy security consideration and limited alternative resource (Ritz, 2014). Asian countries usually put high priority on natural gas security. It means they are less responsive to price movement and this translate into higher willingness-to-pay or equivalently lower price elasticity. Given $\beta_D$ $and$ $\beta_S$, the $\alpha's$ of the consumption and supply equations are calibrated to Japan's monthly consumption data from 2009 to 2017. To maintain the seasonal pattern of the importing countries, the consumption and supply are fitted by month.

As Japan only accounts for ~50% of the entire Asian LNG import, the consumption and supply are scaled by multiplying $\alpha's$ by 2, keeping price elasticity unchanged.

Similar to the U.S., Japan natural gas consumption displays two peaks, one in summer for cooling purpose and the other one in winter for heating purpose. According to Kong (2015), over 50% of natural gas is used for power generation in Japan. China uses most of the natural gas for industrial feedstock and Korea, the second largest natural gas importer use natural gas mainly for power generation as well. Importing regions display similar but not exactly the same pattern compared to the US. On average, U.S. natural gas consumption is about twice the level of shoulder seasons like spring and fall. However, Japan consumption is relatively stable among seasons. The highest month (February) consumes about 30% more natural gas than the lowest month (May).

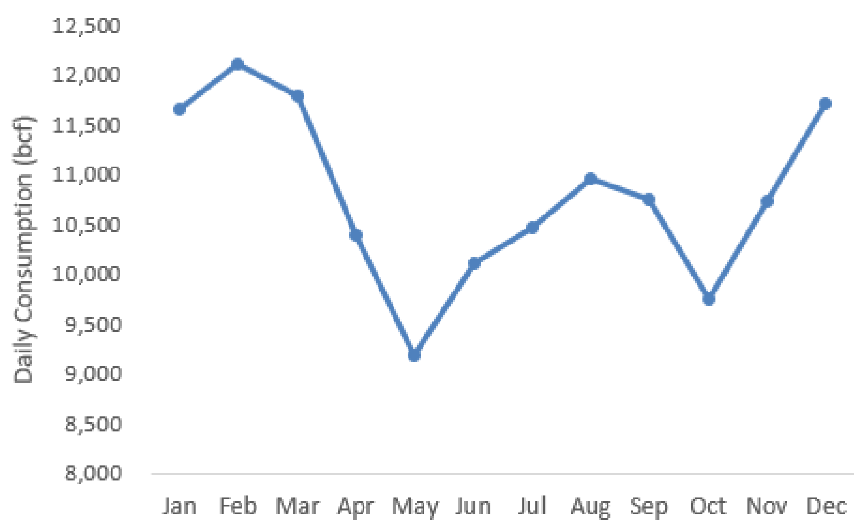


Figure 24. 2009-2017 Japan Monthly Average Daily Consumption (bcf)

Shipping cost is set at $6 per thousand Cubic Feet (Mcf). It is estimated that the liquefaction cost is around $3 per Mcf and shipping rate is about $2 per Mcf from the United States to Asian countries like Japan or China. The regasification process costs about $1 per thousand cubic feet. In total, the cost of whole liquefaction – shipping – re-gas process is commonly estimated to be ~$6 per Mcf to ship natural gas from the United States to Asian countries (Baron, 2014; Medlock, 2012).

In scenario 3, we assume that Asian consumption level remains at current level. The import from other regions of the world like Australia remains at current level as well. This scenario intends to check if there is no demand growth, how much the U.S. exports and the impact will be.

In scenario 4, growth is assumed for both consumption and supply. According to BP Energy Outlook 2017, total natural gas consumption of Asia Pacific area will increase from 631Million tonnes oil equivalent in 2015 by about 100% to 1119 in 2035. Meanwhile, total production of Asia Pacific will only increase from 501 Million tonnes oil equivalent in 2015 by 50% to 756 in 2035. This is consistent with EIA's projection in its Internal Energy Outlook 2013. EIA projects that global natural gas will increase from 21 tcf in 2012 to 45 tcf by 2040, more than doubled. Following this prediction, scenario 4 assumes 100% increase for Asia natural gas consumption and 50% increase for domestic production. The consumption and supply gap is assumed to be filled by the United States LNG export.

We further conduct additional robustness check and sensitivity analysis by changing some of the key parameters, including Asia LNG supply elasticity and demand elasticity. The sensitivity regarding the U.S. domestic supply elasticity is analyzed as well, in scenario 4.5 and 4.6, with increased and decreased price responsiveness respectively.

In order to run simulations for the next 20 years up to 2036, future natural gas production cost and coal price is needed. This study refers to the natural gas production cost estimated by MIT (2010), Paltsev et al (2011) and Deloitte (2012). Future coal price for power generation sector is from EIA's annual Energy Outlook baseline scenario. When the producers extract the natural gas, the sources with cheaper costs is exploited first. The production cost of natural gas will then increase as more natural gas volume is produced over time. MIT (2010), Paltsev et al (2011) and Deloitte (2012) investigate the production cost evolution and find a rough linear correlation between the marginal production cost and the cumulative reserve additions. We use this estimate in our model to represent the production cost as shown in Figure 25: production cost increases linearly along with the cumulative production volume starting from 2016.

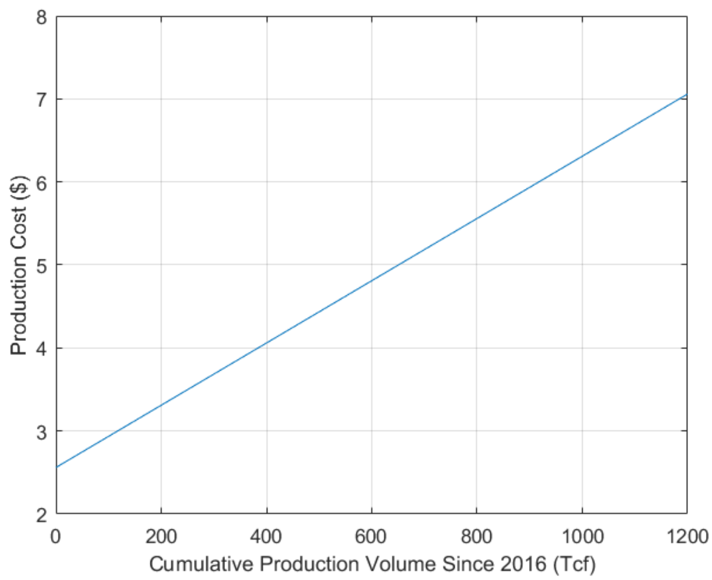

Figure 25. Production cost change with cumulative production

## 6.3 Simulation Results

### 6.3.1 Simulation method

After the competitive storage model with export component is established and the parameters are defined, it can be applied to predict the impact of LNG export on the US domestic natural gas market. The model is solved from January 2010 to December 2036 with the starting values of December 2009. LNG export is assumed to start from June 2017. The first 7 years of simulation is used to make the model stable and a total number of 20 years is simulated after that.

The impact of LNG exports on the US domestic natural gas market is estimated using simulations of the price difference and the structural change with and without export under different scenarios. The simulations are conducted as below:

1. Starting from January 2010, a series of state variables are randomly selected for 324 months (27 years). The regional HDDs and CDDs are randomly selected from the historical observations to keep the regional weather correlation. Consumption shock and production shock are selected as normally distributed random variables.
2. The first set of random variables is applied to the price function of January 2010 to calculate the price ($p_{2010}^{(1)}$). Other variables of the month (consumption, export, storage, and production) can then be determined with the calculated price.
3. The second set of random variables is applied with the calculated prior values of January 2010 to February 2010. Price level ($p_{2010}^{(2)}$) is then calculated along with consumption, export, storage and production, etc.
4. Move forward and continue the simulation for each month until December 2036. A full set of simulated price levels of 20 years is generated with randomly selected state variables.
5. Repeat steps 1-4 with 999 more series of random selected state variables, and a total number of 1000 simulations are collected for price level and standard deviation investigation.

The simulation is applied to each different export scenario including the no export benchmark scenario. The price change due to natural gas export is defined as the difference between the price levels with and without export. If the difference is above zero, it means the export will increase the price level, and vice versa.

6.3.2 LNG Export Impact

As discussed in section 6.2.2, different types of export are applied and compared to investigate the impact of U.S. LNG exports. There are 4 export scenarios and 1 benchmark scenario: scenario (0) is the Benchmark scenario, with no export; scenario (1) is fixed exogenous export of 6 bcf/day; scenario (2) is fixed exogenous export of 12 bcf/day; scenario (3) is endogenous export with fixed demand and supply coefficients ($\alpha_D$ and $\alpha_S$) of the importing countries; scenario (4) is endogenous export with increasing demand and supply coefficients ($\alpha_D$ and $\alpha_S$) of the importing countries. Since all 5 scenarios have no export from 2010 to 2016 and will have same solution with the first 7 years, we focus on the results after January 2017. Figure 26 shows the simulated export quantity associated with all scenarios.

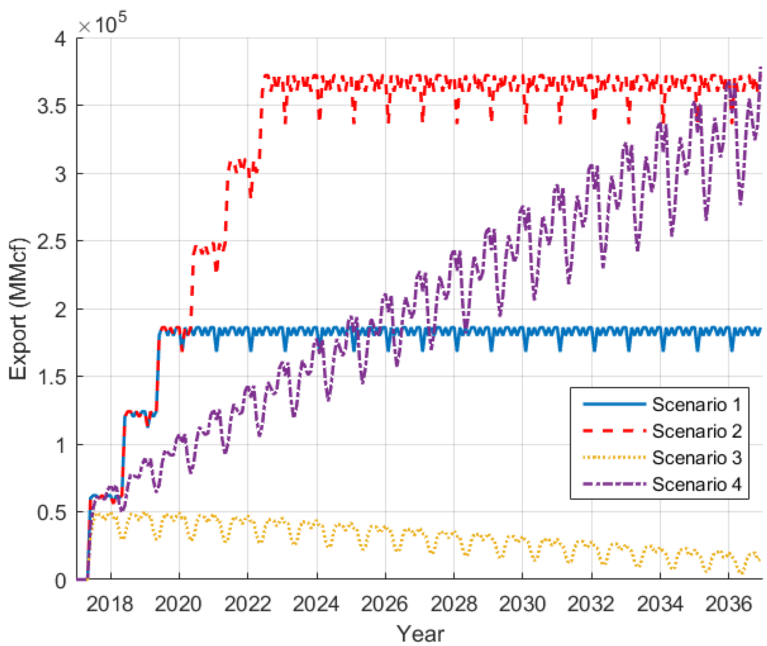


Figure 26. Export from 2017 to 2036 under different scenarios

Figure 26 displays the different export volumes of scenarios 1 – 4. The blue solid line and red dashed line are the export quantities with exogenous scenarios (1 and 2) respectively. The curves follow the expected pattern that the export increases with a steady step (2 bcf/day) every year, until the desired capacities (6 and 12 bcf/day respectively) are reached. The orange dotted and purple dash-dot lines are exports under scenarios 3 and 4 when the export volume is endogenous. With fixed demand and supply coefficients, the importing countries will have the same demand level every year. Due to the increasing cost of production, the US domestic market price increases and accordingly will be less competitive over years compared to the other exporting countries. In Figure 26, we observe a downward-sloping trend for Scenario 3. In Scenario 4, the demand coefficient ($\alpha_D$) increases by 100% and supply coefficient ($\alpha_S$) increases by 50%. The increasing demand cannot be completely fulfilled by the production in the existing importing countries. Sequentially, the import from the US will increase over time.

The simulated monthly price levels in the transition period and in the stable period (last 5 years) of the different scenarios are plotted in Figure 27 and Figure 28 respectively. The price change percentage compared to benchmark scenario are shown in Figure 29. The long-term impact is represented by the results from the last year of simulation (2036), and listed in Table 18 to Table 20. Table 18 shows the values averaged with all months in year 2036, and tables Table 19 and Table 20 represent the values in January and July 2036 respectively to show the seasonal difference.

The thick black solid line in Figure 27 and Figure 28 is the benchmark scenario where no export occurs. Due to the increasing production cost as time goes, the benchmark price with no export increases from \$3 to \$4.8 steadily. The other lines in Figure 27 and Figure 28 represent the price trend with different types of export. The prices start with the same level in 2017. When gradually stabilized in 2036, the price levels generally keep the seasonal pattern but are lifted by a certain amount. With different export scenarios, the transition periods behave differently and last for different length of time.

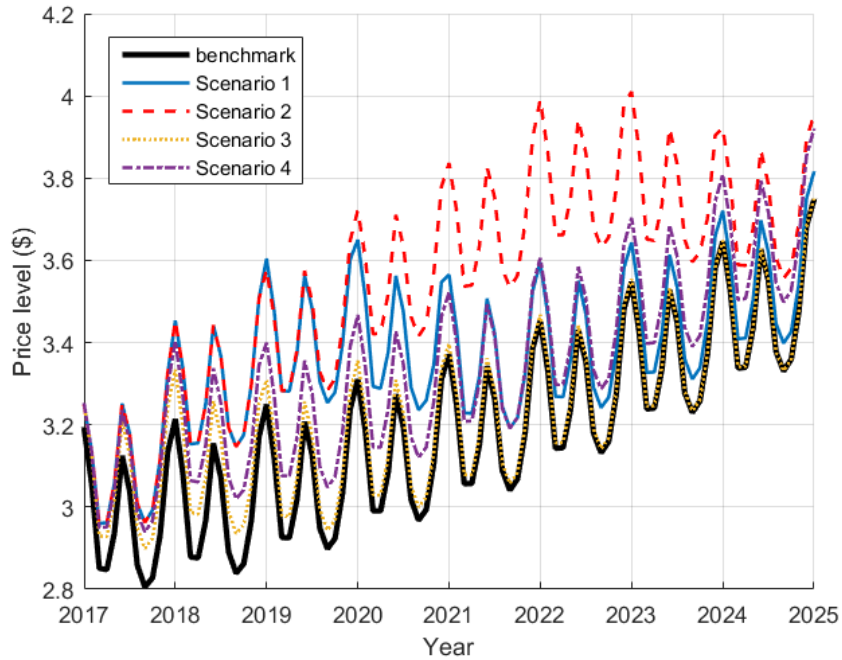


Figure 27. Price level in transition period (2017 – 2025) of simulation

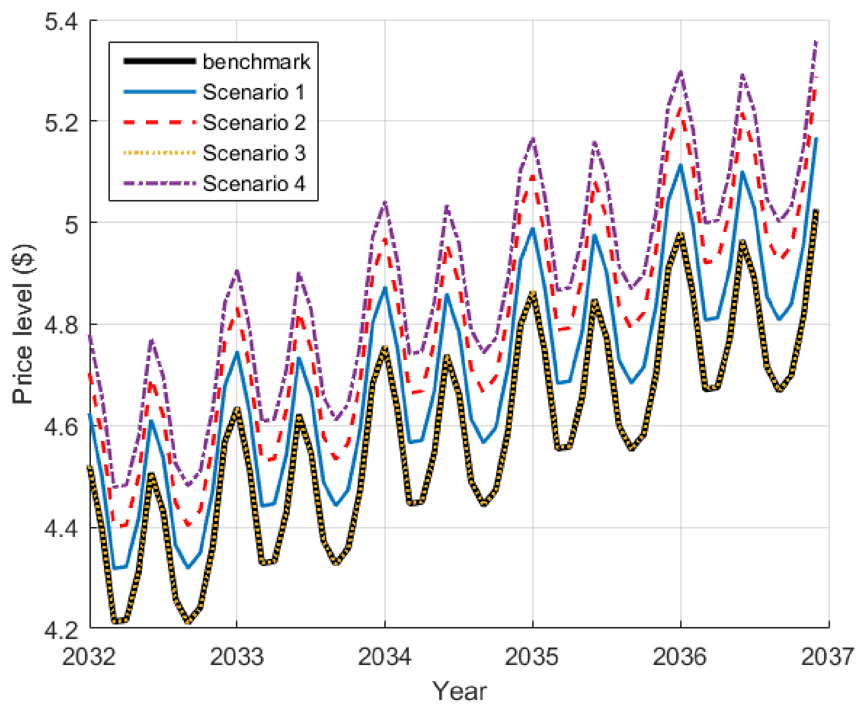


Figure 28. Price level in the last 5 years (2032 – 2036) of simulation

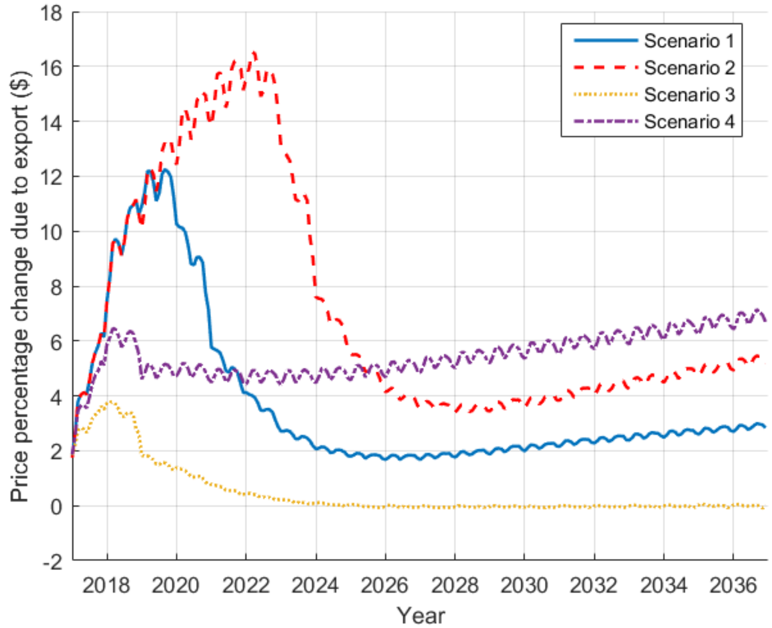


Figure 29. Price change between with and without export

Table 18. Impact on the U.S. domestic market in 2036 – monthly average

| Export Type | Price($) | Price SD | Storage (tcf) | Export (bcf) | Domestic Consumption (tcf) | Production (tcf) |
|---|---|---|---|---|---|---|
| Benchmark | 4.86 | 0.81 | 2.80 | 0 | 1.79 | 2.05 |
| Scenario 1 | 5.01 | 0.85 | 2.80 | 182.5 | 1.78 | 2.21 |
| Scenario 2 | 5.12 | 0.89 | 2.81 | 365 | 1.77 | 2.38 |
| Scenario 3 | 4.86 | 0.74 | 2.80 | 14.537 | 1.79 | 2.06 |
| Scenario 4 | 5.19 | 0.75 | 2.76 | 334.76 | 1.76 | 2.34 |

Table 19. Impact on the U.S. domestic market in January 2036

| Export Type | Price($) | Price SD | Storage (tcf) | Export (bcf) | Domestic Consumption (tcf) | Production (tcf) |
|---|---|---|---|---|---|---|
| Benchmark | 5.10 | 1.01 | 2.55 | 0 | 2.62 | 2.09 |
| Scenario 1 | 5.24 | 1.05 | 2.55 | 186 | 2.60 | 2.26 |
| Scenario 2 | 5.36 | 1.09 | 2.56 | 372 | 2.59 | 2.44 |
| Scenario 3 | 5.09 | 0.92 | 2.54 | 19.692 | 2.62 | 2.11 |
| Scenario 4 | 5.41 | 0.94 | 2.50 | 366.2 | 2.58 | 2.39 |

Table 20. Impact on the U.S. domestic market in July 2036

| Export Type | Price($) | Price SD | Storage (tcf) | Export (bcf) | Domestic Consumption (tcf) | Production (tcf) |
|---|---|---|---|---|---|---|
| Benchmark | 4.91 | 0.70 | 2.88 | 0 | 1.55 | 2.09 |
| Scenario 1 | 5.05 | 0.74 | 2.88 | 186 | 1.53 | 2.26 |
| Scenario 2 | 5.17 | 0.78 | 2.89 | 372 | 1.52 | 2.43 |
| Scenario 3 | 4.91 | 0.64 | 2.90 | 14.417 | 1.55 | 2.10 |
| Scenario 4 | 5.23 | 0.66 | 2.88 | 332.67 | 1.51 | 2.39 |

**Scenario 1 (fixed exogenous export with 6 bcf/day):** The blue solid line in Figure 29 represents the price change with export Scenario 1 compared to no export scenario. The price change experiences a transitional period from 2017 to around 2025, and then moves towards the stabilized state. During the export phase-in period, the price increases rapidly and then it gradually approaches its peak level when the export reaches the 6 bcf per day in 2019. When the export quantity is fixed at 6 bcf/day, natural gas price begins to drop as production catches up with higher demand until 2025. After the transitional period, the price level climbs up gradually with the increase of the production cost.

After year 2025, the US natural gas market moves towards the long-term stable state from the transition period. Natural gas price and volatility increase if there is export. In 2036, the yearly average price increases by 3% due to the LNG export ($4.86 to $5.01), and the price standard deviation increases from $0.81 to $0.85 (16.67% to 16.97% of price level). The same pattern exists for both January and July. In January, the price level increases from $5.10 to $5.24 with price standard deviation increases from $1.01 to $1.05. In July, the price level increases from $4.91 to $5.05 dollars with price standard deviation increases from $0.70 to $0.74.

The price increase results from the lagged domestic production response. If the production increase cannot catch up quickly enough to satisfy the LNG export, price will increase and domestic consumption will decrease. The production catch-up refers to the difference of production level with and without export. Figure 30 shows the production change with the export for Scenario 1, and it describes how much and how fast production catches up with exports. When production stabilizes after 2025, there still exists a gap between the amount of production increase and the fixed export level. As the export level is a fixed quantity, the natural gas availability to the U.S. customers is squeezed to a lower level, and the price is more sensitive to shocks and experiences a higher variance.

From 2017 to 2025, the US natural gas market experiences the transition stage due to LNG exports. As exports increase with a 2 bcf/day per year rate from 2017 to 2019, price increases also and approaches the peak level of 12% higher in 2019. After 2019, exports are stable while production is catching up rapidly due to higher price compared to production cost (Figure 30). As a result, natural gas price begins to drop, and a 6% price drop is observed in the first 2 years (2020 – 2021). From 2022 to 2025, the catch-up rate of production slows down, and the price change gradually drops to 2%, which ends the transition stage.

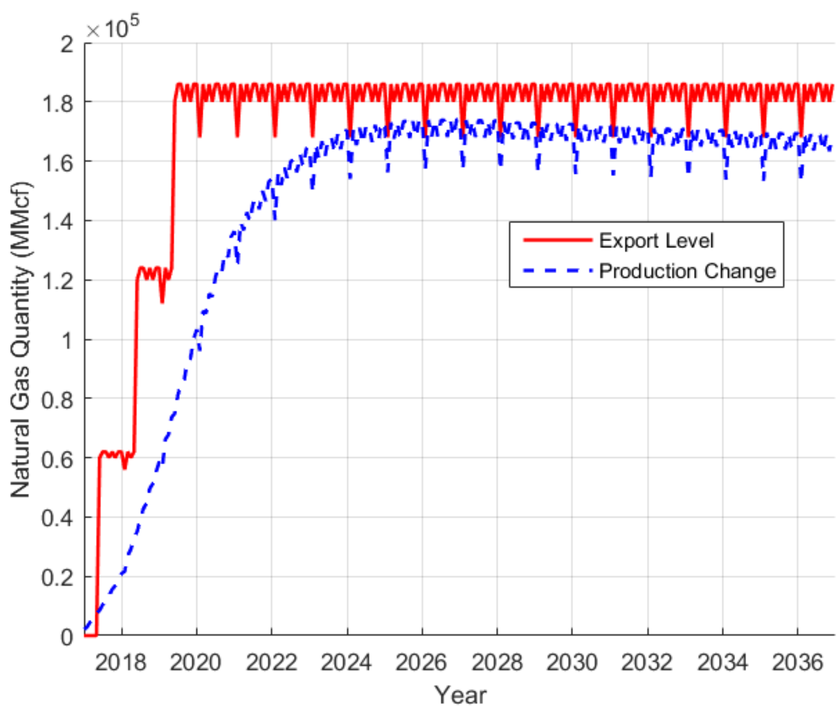


Figure 30. Production change vs. export level in Scenario 1

When LNG exports are introduced, considering that the storage change is constrained by the facility capacity, exports are fulfilled either by increased production or the decreased domestic consumption or inventory. How much export is covered by each aspect is of interest and describes how the US domestic natural gas market is affected. Figure 31 shows the percentage of export absorbed by production change and consumption change. When exports are first introduced to the market, production cannot respond quickly enough, and more than 70% of exports come from reduced consumption and about 25% of exports come from the production increase. Inventory drawdown satisfies the remaining small portion. As the production level increases gradually, the percentage of export covered by production increases, and the US domestic consumption begins to recover. From 2017 to 2025, the percentage covered by production increases from 30% to more than 90%, and the US natural gas market also recovers to a new stable state.

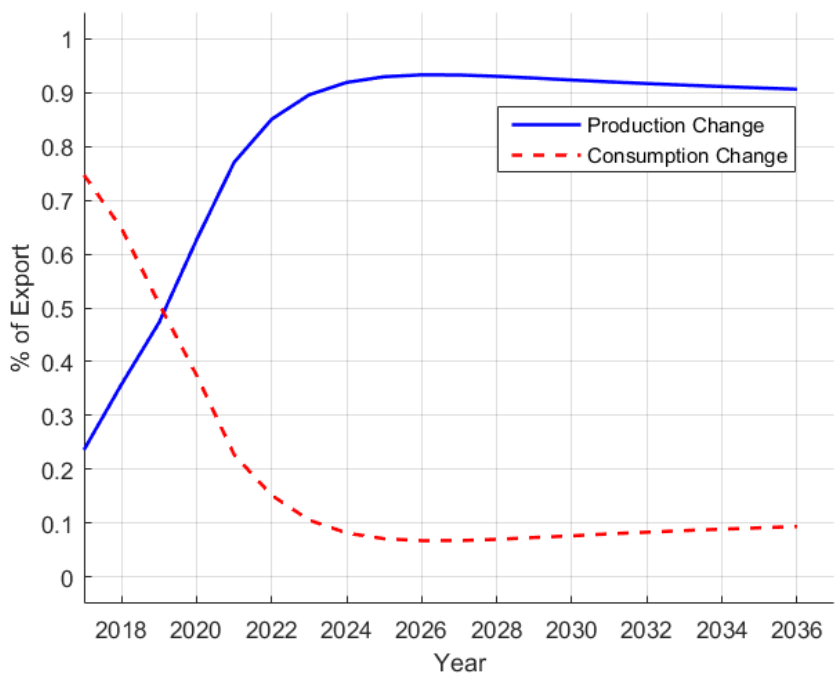


Figure 31. Percent of export covered by production and consumption changes in Scenario 1

**Scenario 2 (fixed exogenous export with 12 bcf/day):** The red dashed line in Figure 29 represents the price change of Scenario 2. The price change curve has a shape similar to that of Scenario 1. The price change increases as exports increase from 2017 to 2022, reaching a higher peak as the export volume is larger. After that the price drops for the following 5 years up to 2027. It then increases slowly again as the production cost increases.

Similar to Scenario 1, the US natural gas market moves into a stable stage after 2027. With a higher fixed export level (12 bcf/day), the natural gas price level and price standard deviation increase to higher values. In 2036, the average price is higher by 5.3% from $4.86 to $5.12 and the price standard deviation increases from $0.81 to $0.89. In January the price increases from $5.10 to $5.36 with price standard deviation increases from $1.01 to $1.09. In July the price increases from $4.91 to $5.17 with price standard deviation increased from $0.70 to $0.78.

The production catch-up pattern of Scenario 2 is shown in Figure 32. Production increases catch up with exports rapidly in the beginning, and slows down after 2027. In the stable stage, the production increase is less than the LNG export level. As discussed in Scenario 1, this will lead to a reduction of natural gas supplied to domestic market and more sensitive market price. That is, the price level is elevated and price standard deviation also increases.

Compared to Scenario 1, Scenario 2 has a similar pattern of transition stage while with a longer period (2017 – 2027). The price change climbs up from 2017 and reaches its peak value of 16% in 2022 when export approaches the capacity. After the export is fixed at the defined level, the production still increases at high speed for the following 2 years from 2023 to 2024, and the price change drops by rapidly in these 2 years. From 2025 to 2027, the production increases at a slower rate, and the price increase accordingly drops gradually to 4%.

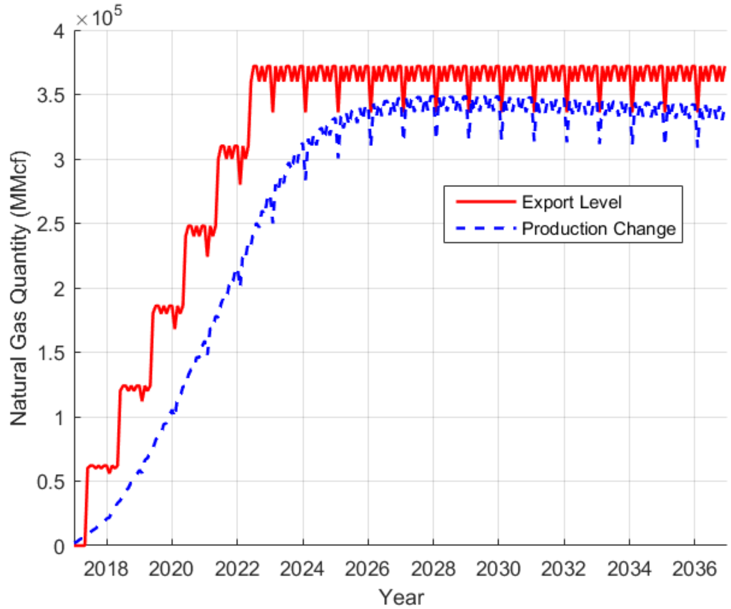


Figure 32. Production change vs. export level in Scenario 2

Similar to the changes in Scenario 1, the consumption decrease covers more than 70% of the export level in the beginning, and production increase only covers 25%. In the transitionary period, the production keeps increasing and the consumption steadily recovers as the production catches up with exports. In the long-term stable stage, more than 90% of export is provided by production, and consumption reduction only accounts for less than 10%.

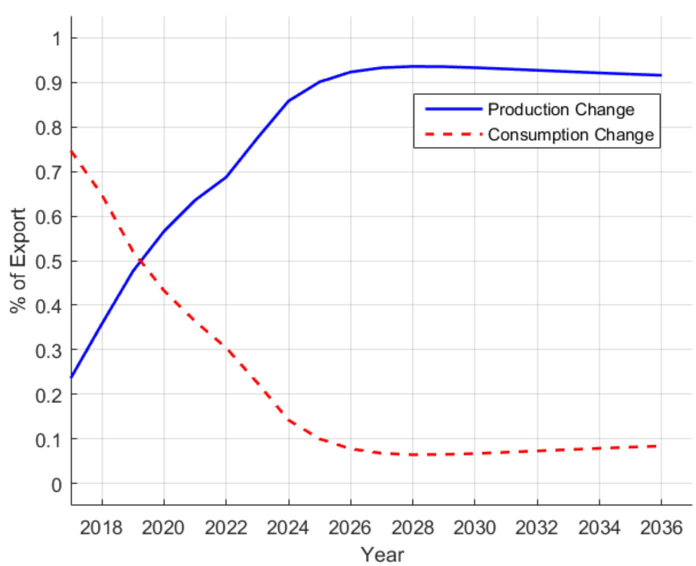


Figure 33. Percent of export covered by production and consumption changes in scenario 2

**Scenario 3 (endogenous export without growth):** The dotted orange line in Figure 29 represents the price change under Scenario 3. With endogenous model assumption, the export level displays a clear seasonal pattern and is a function of the US market price. The price change increases following the export increase in 2017, and drops after that due to export reduction and production increase. Beginning in 2025, the price drops to the same level as when there is no export, and then stabilize for the rest of the simulation horizon.

In the stable stage after 2025, the average price is similar to the benchmark scenario while with a smaller price standard deviation. Compared to benchmark results, the yearly average price of Scenario 3 is at the same level ($4.86), but the price standard deviation reduces from $0.81 to $0.74. The prices of different seasons behave similarly. In January, the price is almost the same ($5.10 in benchmark scenario while $5.09 under scenario 3) with price standard deviation reduced from 1.01 to 0.92 dollar. In July, the price level is also the same at $4.91 with price standard deviation reduces from $0.70 to $0.64.

In Scenario 3, the Asian natural gas demand and supply coefficients are assumed to be constant and do not change with time. As the US market price of natural gas keeps increasing with production cost, the US will become less competitive and export less as time goes. Figure 34 shows the export level and production change with time. The export level is low in Scenario 3. Compared to the US consumption level, the export in 2018 equals to only 2.1% of US total consumption of the same year. As the export level is low and has a downward sloping trend, it is easily met by the production increase as shown in Figure 34. With the export completely covered by the increased production, the average US consumption and price will also recover to the level close to the benchmark scenario.

The impact on the price variance from the endogenous LNG export is also of interest. With the export level negatively correlated with natural gas price, the price becomes less volatile. The export can act as a buffering sector to the US market. When the US market experiences unexpected shock which tends to increase the price level, exports fall due to higher price. The reduced export will be consumed in the US market. In this manner, the price impact of the shock will be reduced.

The transition period of Scenario 3 is from year 2017 to 2025. The price increases by 4% in the beginning when export is initially introduced. After that the price keeps dropping as exports decrease and production catches up rapidly. In 2025 the production completely catches up with the export level and the price change also drops to zero.

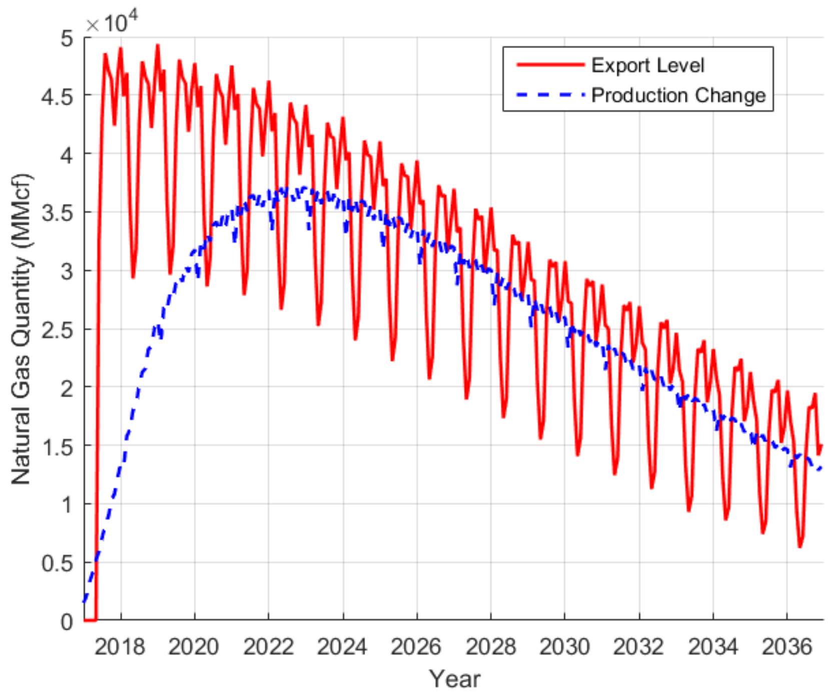


Figure 34. Production change vs. export level in Scenario 3

The production and consumption changes as a percentage of the export level are plotted in Figure 35. The production change begins with a low percentage of 25% and catches up to a complete 100% coverage of the export level. We can see that the production change after 2026 actually covers slightly more than 100%. This means that due to the reduction of exports the production increase overshoots a small amount when catching up, and this also explains why the price level can sometimes be lower than the benchmark scenario.

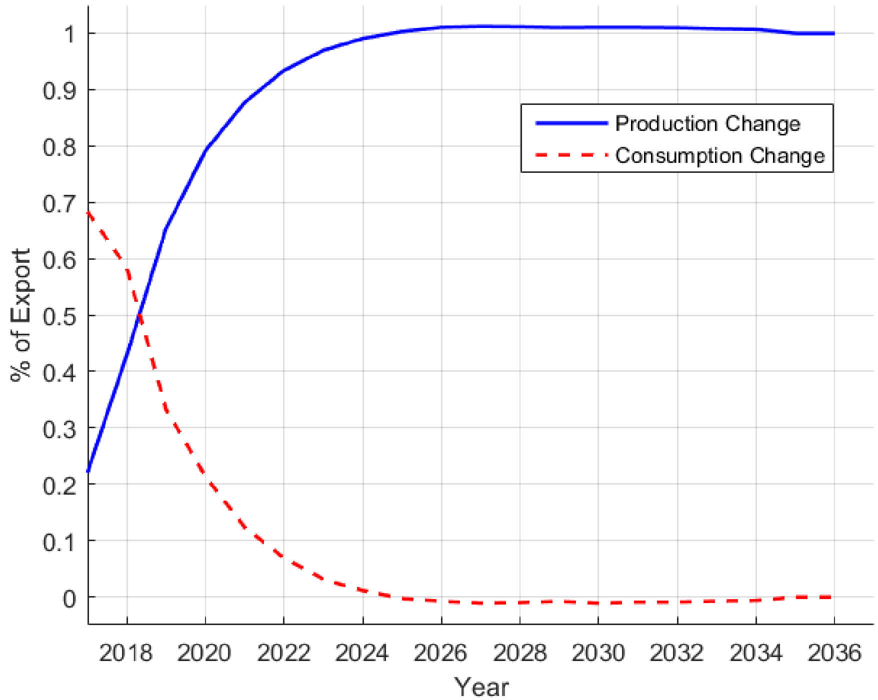


Figure 35. Percent of export covered by production and consumption changes in scenario 3

**<u>Scenario 4 (endogenous export with growth):</u>** The purple dash-dot line in Figure 29 represents the price change under scenario 4. Unlike Scenario 3, due to the increased Asian natural gas demand and supply, the export level keeps increasing from 2017 to 2036. Accordingly, after an initial jump, the price change increases steadily beginning in 2019.

As the export increases steadily after the first year of introduction, the price change increases gradually from 2019. In 2036, compared to the benchmark scenario with no natural gas export, the average yearly price level increases by 6.8% from \$4.86 to \$5.19, and the price standard deviation reduces from \$0.81 to \$0.75. The seasonal price acts similarly. In January, the price increases from \$5.10 to \$5.41 with price standard deviation reduced from \$1.01 to \$0.94. In

July, the price increases from \$4.91 to \$5.23 with price standard deviation reduces from \$0.70 to \$0.66.

Higher LNG export volumes lead to higher prices. In 2017, the export level is low and the price change is not significant. As a result, the production increase is not as large as in Scenario 1 and 2. Figure 36 shows that the production change keeps increasing but at a lower level compared with the export increase. In this manner, the production available to the US domestic market is constantly lower. The price then keeps increasing till the end of the simulation horizon. The price variance reduces due to the buffering effects from the endogenous export as discussed in Scenario 3. The price is less volatile with the negative-correlated export function of price.

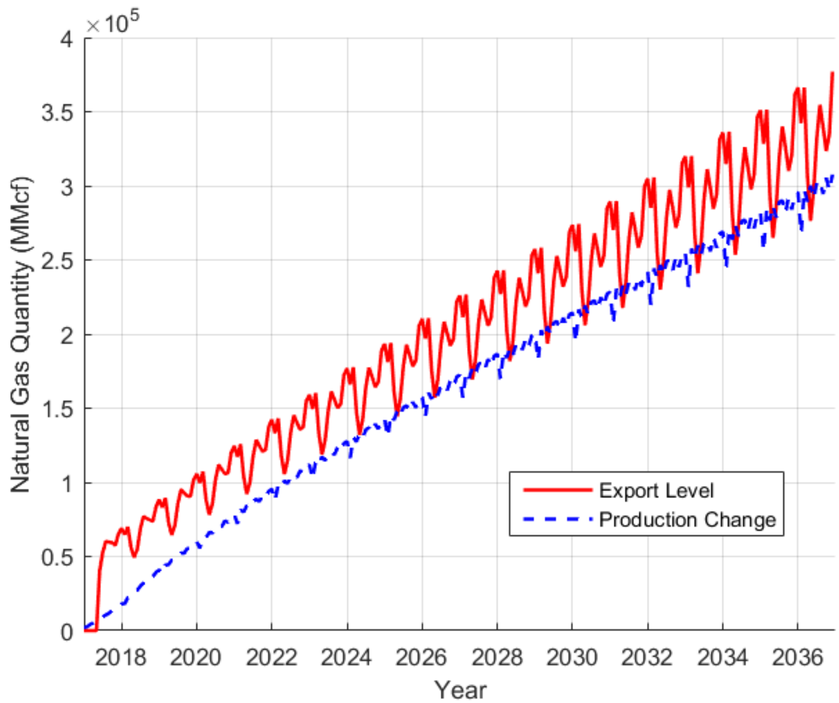


Figure 36. Production change vs. export level in Scenario 4

As shown in Figure 37, in the beginning periods 25% of the export comes from production increase, and 70% comes from consumption reduction. The remaining 5% comes from inventory drawdown. With the increasing export level, the production catches up at a slower pace. It covers less than 90% of the total LNG export until 2036. This leads to a lower supply level in US domestic market, and explains why Scenario 4 has the highest price increase in the stationary stage.

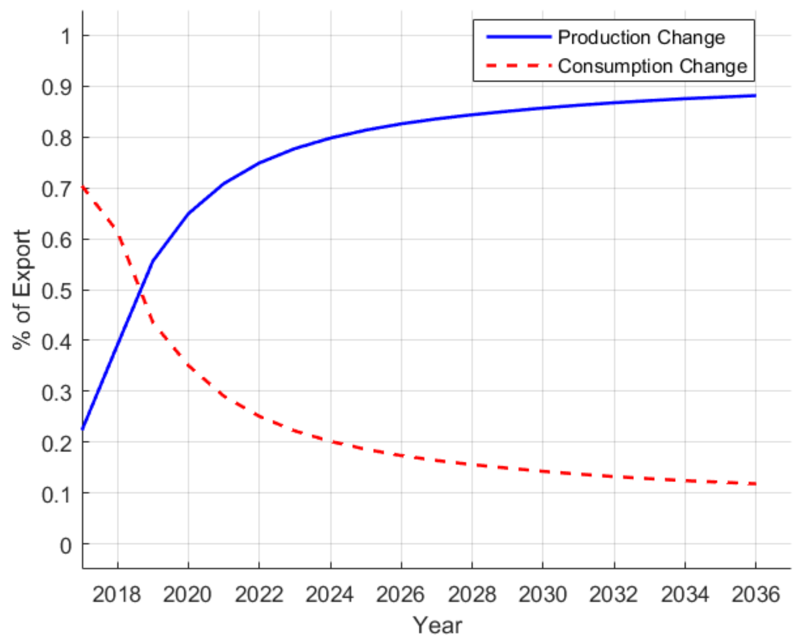


Figure 37. Percent of exports covered by production and consumption changes in scenario 4

6.3.3 Robustness check and Sensitivity Analysis

For endogenous LNG export scenarios, the results might change if some of the key parameters are specified differently. This study conducts additional scenario simulations to check the robustness of the model results. In addition, scenarios using different model parameters shed some light on the result sensitivity to key parameters. The sensitivity of three key parameters are checked: importing countries' price elasticity of LNG demand, price elasticity of LNG supply from regions other than the United States and the U.S. domestic production price responsiveness.

The results are overall consistent with the scenarios discussed above in the long-term. The short-term variation is relatively larger if using different parameters.

LNG Supply Elasticity

The price elasticity of Asian LNG supply from other regions other than the United States is set at 0.1. The supply elasticity may be higher or lower than this number. The robustness check is based on scenario 4 with growth assumption. In Scenario 4.1, the supply elasticity is decreased by 50% to 0.05 while the elasticity is doubled in Scenario 4.2 to 0.2. The long-term impact is small and consistent with scenario 4 as we can see from Table 21.

Table 21. Robustness Check regarding Price Elasticity of LNG Supply from Other Regions

| | Scenario | Price ($) | Price SD | Stor. (tcf) | Export (bcf) | Domestic Cons. (tcf) | Prod. (tcf) |
|---|---|---|---|---|---|---|---|
| | Benchmark (Scenario 4) | 5.41 | 0.94 | 2.50 | 366.20 | 2.58 | 2.389 |
| | 0.5* LNG supply elasticity | 5.42 | 0.99 | 2.50 | 359.40 | 2.58 | 2.385 |
| | 2* LNG supply elasticity | 5.39 | 0.86 | 2.49 | 377.01 | 2.58 | 2.396 |
| | Benchmark (Scenario 4) | 5.23 | 0.66 | 2.88 | 332.67 | 1.51 | 2.390 |
| | 0.5* LNG supply elasticity | 5.24 | 0.70 | 2.87 | 328.54 | 1.51 | 2.386 |
| | 2* LNG supply elasticity | 5.22 | 0.60 | 2.90 | 337.01 | 1.51 | 2.396 |
| | Benchmark (Scenario 4) | 5.19 | 0.75 | 2.76 | 334.76 | 1.76 | 2.345 |
| | 0.5* LNG supply elasticity | 5.20 | 0.79 | 2.76 | 330.65 | 1.76 | 2.341 |
| | 2* LNG supply elasticity | 5.17 | 0.68 | 2.77 | 339.91 | 1.76 | 2.351 |

In the transition period, natural gas price of the U.S. increases as production is gradually increased to satisfy LNG export. Asian LNG price increases accordingly as its price equals the U.S. price plus $6 shipping cost when exports exist. Asian imports from other regions decrease along with the increased price but to a lesser extent due to a smaller supply elasticity. As a result, the export needed from the U.S. decreases and the U.S. domestic price decreases due to lower total demand, as illustrated in Figure 38. In the long term, as price gradually reverts back, the impact diminishes and the price impact is very close to 0. The same logic applies when supply elasticity increases.

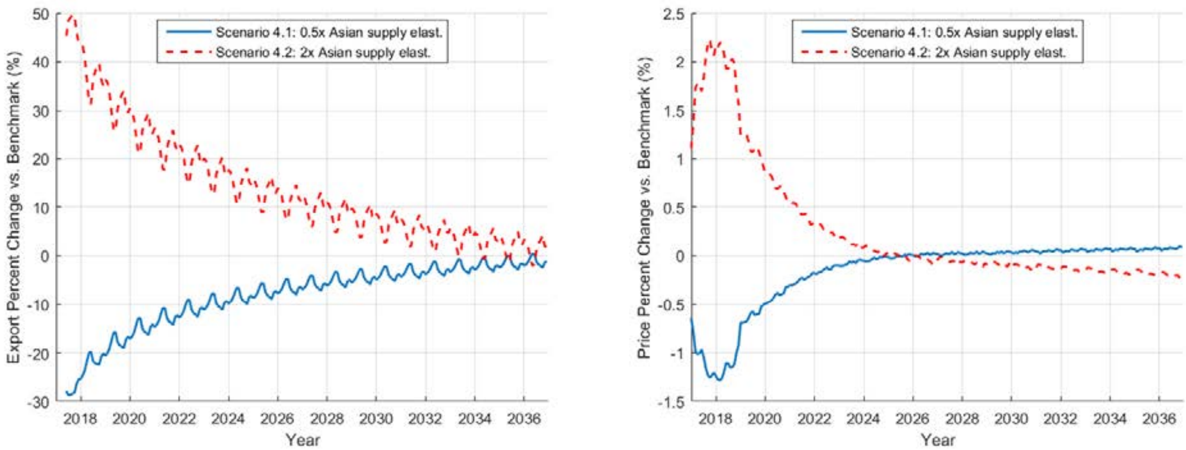


Figure 38. Export and Price Change due to Alternative LNG Supply Elasticity

<u>Demand Price Elasticity</u>

The current Asian natural gas demand elasticity is set at -0.06. This number may increase in the future as they diversify the energy sources and more alternative fuels are available. We consider the scenario when demand price elasticity increases. In Scenario 4.3, the demand elasticity increases by 100% to -0.12. The elasticity triples in Scenario 4.4 from -0.06 to -0.18. The long-term impact is small and consistent with scenario 4 as we can see from Table 22. When Asian demand elasticity increases, U.S. price standard deviation decreases. Higher demand elasticity means users are more responsive to price change. When price are high, import decrease more with higher demand elasticity. The decreased import helps reduce high price pressure and thus variance decreases.

In the first several years right after U.S. LNG exports begin, exports increase compared to scenario 4 if demand elasticity is higher. The availability of the U.S. natural gas decreases the Asian price. If demand is more sensitive to price change, Asia imports will increase more due to the price decrease. The higher the demand elasticity, the larger the import will be. Higher import pushes domestic natural gas price up. The impact fades away when price is getting close to starting level.

Table 22. Robustness Check regarding Asian Demand Elasticity

| | Scenario | Price ($) | Price SD | Stor. (tcf) | Export (bcf) | Domestic Cons. (tcf) | Prod. (tcf) |
|---|---|---|---|---|---|---|---|
| | Benchmark (Scenario 4) | 5.41 | 0.94 | 2.50 | 366.20 | 2.58 | 2.389 |
| | 2* Asian demand elas. | 5.41 | 0.91 | 2.49 | 373.79 | 2.58 | 2.395 |
| | 3* Asian demand elas. | 5.40 | 0.88 | 2.49 | 381.93 | 2.58 | 2.402 |
| | Benchmark (Scenario 4) | 5.23 | 0.66 | 2.88 | 332.67 | 1.51 | 2.390 |
| | 2* Asian demand elas. | 5.23 | 0.63 | 2.89 | 338.66 | 1.51 | 2.396 |
| | 3* Asian demand elas. | 5.23 | 0.61 | 2.90 | 345.33 | 1.51 | 2.402 |
| | Benchmark (Scenario 4) | 5.19 | 0.75 | 2.76 | 334.76 | 1.76 | 2.345 |
| | 2* Asian demand elas. | 5.18 | 0.72 | 2.77 | 340.27 | 1.76 | 2.350 |
| | 3* Asian demand elas. | 5.18 | 0.70 | 2.77 | 346.35 | 1.76 | 2.356 |

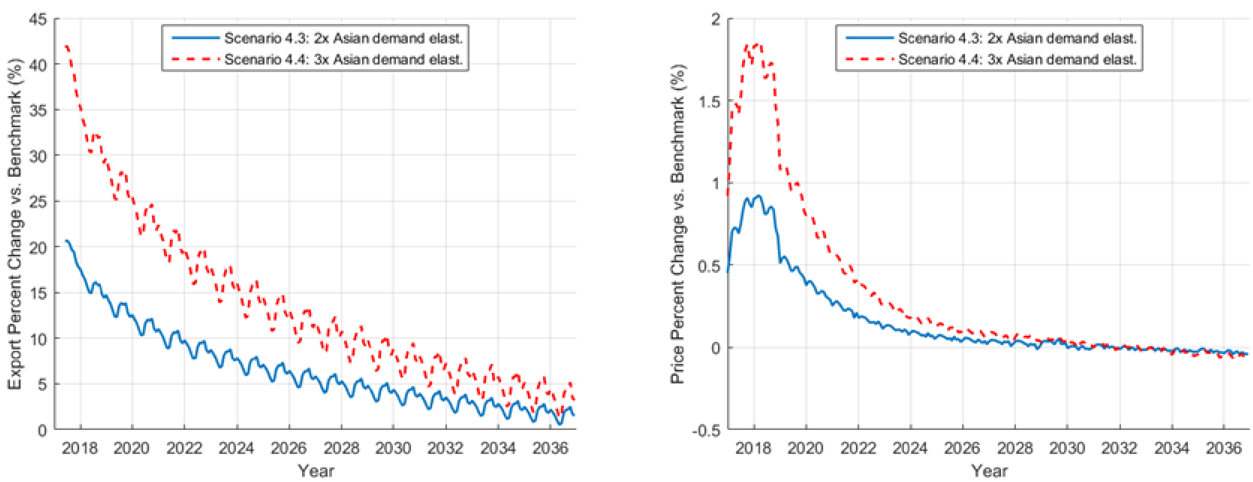


Figure 39. Export and Price Change due to Alternative Asian Demand Elasticity

<u>Sensitivity to domestic production elasticity</u>

Both the long-term and transitionary period price impact is highly dependent on how fast U.S. domestic producers can increase their production capacity and meet the extra natural gas demand resulting from LNG export. The U.S. production function is specified as below with estimated investment incentive parameter $\beta_1$ at 0.014 and estimated short-term price elasticity at 0.008.

$$Pr^{(t,m)} = \beta_0 Pr^{(t,m-1)} \left(\frac{E_{t-1}(P_{mean}^{(t)})}{cost^{(t)}}\right)^{\beta_1} \left(\frac{E_{t,m}(p^{(t,m+1)})}{E_{t-1}(P_{mean}^{(t)})}\right)^{\beta_2} \varepsilon_{pr}$$

If the investment incentive parameter decreases by 50% from 0.014 to 0.007, the speed of the production catch-up to meet LNG export will be much slower. In addition, higher price is needed to produce the same amount of natural gas in the long term. This will translate into higher natural gas prices in both the short-term and long term. Comparing to scenario 4.0 that uses the calibrated production function, price is \$0.50 or around 10% higher by 2036. Consequently, exports will decrease by about 4% in the long term.

On the other side, if the investment incentive parameter of production function doubled from 0.014 to 0.028, producers are more sensitive to price increase. More natural gas will be produced and the price decreases compared to scenario 4.0. The price decrease is more pronounced in the transitionary period due to tight supply during that period. The LNG export volume increase by above 2% in the long term.

In terms of price variance, price have larger variance when production is less sensitive to price change. The price standard deviation is 0.75 in benchmark scenario 4.0. When production becomes more responsive to price change, price standard deviation decreased to 0.72. When production is less responsive to price change, price stand deviation increased to 0.81 end of simulation horizon, year 2036. More details can be seen in Table 23.

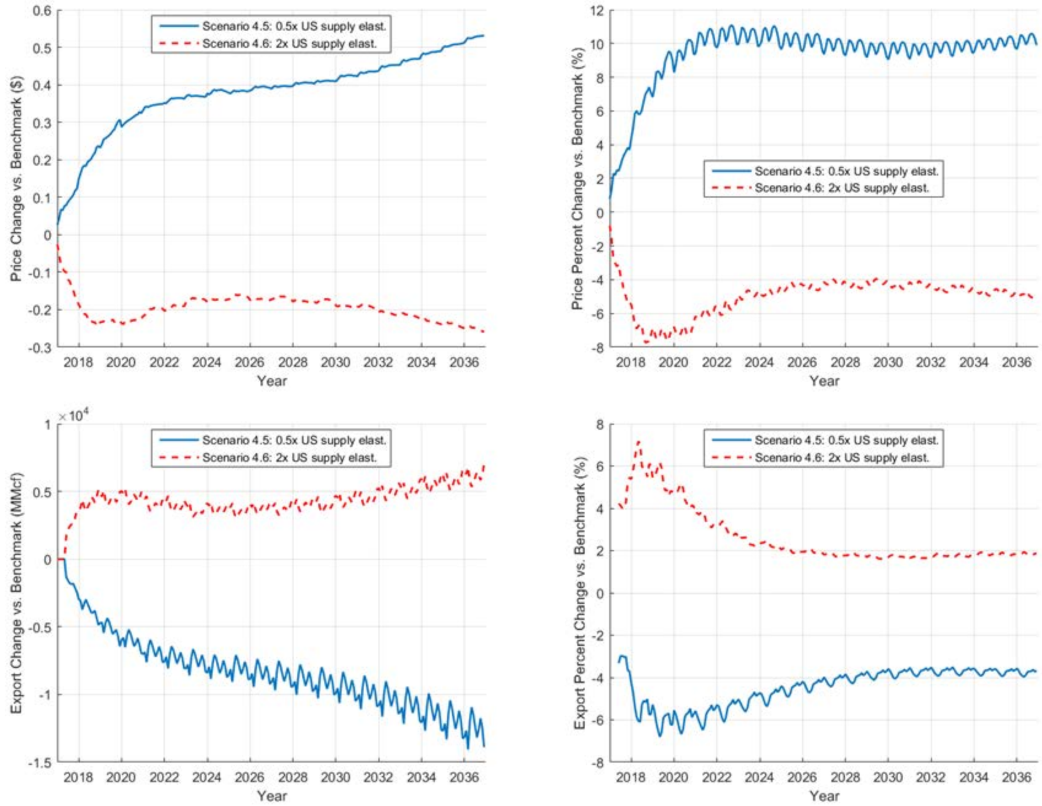


Figure 40. Price and Export Comparison under Different Production Specification

Table 23. Sensitivity to the U.S. Production Elasticity

| | Scenario | Price ($) | Price SD | Stor. (tcf) | Export (bcf) | Domestic Cons. (tcf) | Prod. (tcf) |
|---|---|---|---|---|---|---|---|
| | Benchmark (Scenario 4) | 5.41 | 0.94 | 2.50 | 366.20 | 2.58 | 2.389 |
| | 0.5* US supply elasticity | 5.94 | 1.00 | 2.50 | 352.84 | 2.52 | 2.320 |
| | 2* US supply elasticity | 5.15 | 0.91 | 2.49 | 373.17 | 2.61 | 2.427 |
| | Benchmark (Scenario 4) | 5.23 | 0.66 | 2.88 | 332.67 | 1.51 | 2.390 |
| | 0.5* US supply elasticity | 5.77 | 0.72 | 2.88 | 320.17 | 1.46 | 2.320 |
| | 2* US supply elasticity | 4.98 | 0.63 | 2.88 | 338.93 | 1.54 | 2.428 |
| | Benchmark (Scenario 4) | 5.19 | 0.75 | 2.76 | 334.76 | 1.76 | 2.345 |
| | 0.5* US supply elasticity | 5.72 | 0.81 | 2.77 | 322.08 | 1.71 | 2.276 |
| | 2* US supply elasticity | 4.92 | 0.72 | 2.76 | 341.18 | 1.78 | 2.382 |

# CHAPTER 7. DISCUSSION AND CONCLUSION

This thesis aims to construct a model that can mimic the major market participants' behavior and reproduce natural gas prices with major seasonal patterns consistent with historical observations. A monthly rational-expectations competitive storage model is constructed to better reflect monthly variations in price. The model is based on market fundamentals to facilitate simulations of natural gas demand shifts or changes in production technologies.

Back-testing results from 2010 to 2016 show that the model can simulate price series that are reasonably close to historical observations. The model generates simulated prices that largely replicate the key features of historical data, including the price level, price variance, price sensitivity under unusual weather conditions and high price autocorrelation. Weather conditions and total natural gas availability are the main drivers for price and price standard deviation. The model finds that in winter, high heating degree days (HDD) or low inventory drives price and price volatility higher while price and its variance decrease with low HDD and high inventory. The scenario is similar in summer with cooling degree days (CDD) instead of HDD as the weather variable. When inventory is low, weather shocks have a larger impact on price than when inventory is high. The effect is more pronounced in winter than in summer because the supply is tighter in heating season.

The model is solved using numerical methods because analytical solutions are not feasible. In order to balance accuracy and meanwhile achieve reasonable running time, natural gas total availability, regional weather variables and total consumption shock are included as state variables. To facilitate model solving, we make two assumptions: 1) using mean weather variables HDD/CDD without shock to approximate the prior consumption and production level to feed into next month function; 2) the coefficients of the price function of January in adjacent years are the same. Accuracy tests show that the model's approximation errors are reasonable. The model may perform better if we were able to include prior consumption and production level as state variables. However, "curse of dimensionality" exists and the inclusion of additional state variables will make the computation time prohibitively long.

The monthly rational-expectations storage model is useful for analyzing the LNG export impact because the inclusion of storage makes the model dynamic and prices become serially correlated so that the model can capture both short-term and long run market response. In addition, as foreign countries' gas consumption patterns throughout one year are different from the U.S., the impact of gas export for each month or different season might be different. The utilization of a monthly model is able to assess the impact on each single month. This study covers some of the very aggressive export scenarios like 12 bcf per day and U.S. LNG is used to fill the Asian consumption and supply growth gap. LNG export impact is assessed under different export contract types. The exogenous scenario assumes a fixed export volume for all periods and no seasonal pattern is considered. This is the case when imported U.S. natural gas is mostly used for baseload consumption and long-term contracts are signed. When the global natural gas market gradually transitions from isolated to integrated market, competitive international trade framework can be utilized to evaluate the LNG impact on the U.S. domestic market. Scenarios 3 and 4 assume the law of one price holds and export volumes are endogenously determined. Below are the major findings of this study.

With high shipping cost and inelastic Asian natural gas demand, U.S. LNG is not competitive under current market condition if no Asian demand growth is expected. As shown in scenario 3, if Asia natural gas consumption and import from other regions remains at current level, the U.S. LNG export volume is very small and decreases over time. Due to small export volumes, the domestic price impact is minimal.

The long-term price impact is less than 8% under all scenarios analyzed in this study, or around $0.33 per thousand cubic feet. In 2036, the endogenous export with growth assumptions scenario shows the largest price increase compared to the no export benchmark scenario. The export level is around 12 bcf per day. The price impact is small in general due to relative small volume: total US domestic consumption level is 27 tcf in 2016, equal to 74 bcf per day. If the US is going to export 12 bcf/d, this accounts for about 16% of the US total consumption. If the export volume is set at 6 bcf per day, it will only account for 8% of the total U.S consumption.

Price volatility becomes smaller if an endogenous export sector is added while the price volatility becomes higher under fixed export volume scenarios. If the LNG export is endogenously determined, when domestic price increased, LNG export decreases. This provides an additional price buffer if there is U.S. domestic demand shock that increases natural gas consumption. On the contrary, fixed volume exports makes the total natural gas consumption less responsive to price changes and thus increases price variance.

Most of the LNG export volumes will be met by production increases rather than domestic consumption decreases in the long term. In all four scenarios analyzed in this study, production catches up gradually in response to a price increase due to LNG export. In the beginning period when production is constrained by production capacity, most of the export is covered by domestic consumption reduction. In the long term, as production increase, domestic consumption recovers to similar level as in the no export scenario.

REFERENCES

Alterman, S., 2012. Natural gas price volatility in the UK and North America. Oxford Institute for Energy Studies, ISBN 978-1-907555-43-5

Arora, V., 2014. Estimates of the price elasticities of natural gas supply and consumption in the United States. *MPRA Paper* No. 54232.

Baron, R., Bernstein, P., Montgomery, W.D. and Tuladhar, S., 2015. Macroeconomic impacts of LNG exports from the United States. *Economics of Energy & Environmental Policy, 4*(1).

Bernstein, P., Tuladhar, S.D. and Yuan, M., 2016. Economics of US natural gas exports: Should regulators limit US LNG exports?. *Energy Economics*, *60*, pp.427-437.

Boogert, A. and De Jong, C., 2008. Gas storage valuation using a Monte Carlo method. *The Journal of Derivatives*, 15(3), pp.81-98.

BP Energy Outlook 2035, 2014. Available from: *http://www.bp.com/content/dam/bp/pdf/energy-economics/energy-outlook-2016/bp-energy-outlook-2014.pdf*

BP Energy Outlook 2017, 2017. Available from: *http://www.bp.com/en/global/corporate/energy-economics/energy-outlook/energy-outlook-downloads.html*

Cafiero, C. and Wright, B.D., 2011. The empirical relevance of the competitive storage model. *Journal of Econometrics*, *162*(1), pp.44-54.

Carmona, R. and Ludkovski, M., 2010. Valuation of energy storage: An optimal switching approach. *Quantitative Finance*, 10(4), pp.359-374.

Chambers, M.J. and Bailey, R.E., 1996. A theory of commodity price fluctuations. *Journal of Political Economy*, *104*(5), pp.924-957.

Chiou-Wei, S.Z., Linn, S.C. and Zhu, Z., 2014. The response of US natural gas futures and spot prices to storage change surprises: Fundamental information and the effect of escalating physical gas production. *Journal of International Money and Finance*, 42, pp.156-173.

Dahl, C.A. and Roman, C., 2004. A Survey of Energy Demand Elasticities. IAEE Working P Cleveland, OH: International Association of Energy Economics.

Dahl, C.A. and Sterner, T., 1991. Analysing gasoline consumption elasticities: a survey. *Energy economics*, *13*(3), pp.203-210.

Davis, L.W. and Muehlegger, E., 2010. Do Americans consume too little natural gas? An empirical test of marginal cost pricing. *The RAND Journal of Economics*, 41(4), pp.791-810.

De Jong, C. and Walet, K., 2003. To store or not to store. *Energy Risk Magazine* (October): S8-S11.

De Joode, J. and Özdemir, Ö., 2010. Consumption for seasonal gas storage in northwest Europe until 2030: Simulation results with a dynamic model. *Energy Policy*, 38(10), pp.5817-5829.

Deaton, A. and Laroque, G., 1992. On the behaviour of commodity prices. *The Review of Economic Studies*, *59*(1), pp.1-23.

Deaton, A. and Laroque, G., 1995. Estimating a nonlinear rational expectations commodity price model with unobservable state variables. *Journal of Applied Econometrics*, *10*(S1).

Deaton, A. and Laroque, G., 1996. Competitive storage and commodity price dynamics. *Journal of Political Economy*, *104*(5), pp.896-923.

Deaton, A. and Laroque, G., 2003. A model of commodity prices after Sir Arthur Lewis. *Journal of Development Economics*, *71*(2), pp.289-310.

Deloitte Center for Energy Solutions, 2011. Made in America, the economic impact of LNG exports from the United States. https://www2.deloitte.com/us/en/pages/energy-and-resources/articles/made-in-america-the-economic-impact-of-lng-exports-from-the-united-states.html

Dubin, J.A. and McFadden, D.L., 1984. An econometric analysis of residential electric appliance holdings and consumption. *Econometrica: Journal of the Econometric Society*, pp.345-362.

EIA, 2014. Effect of increased levels of Liquefied Natural Gas exports on U.S. energy markets. Available from: *https://www.eia.gov/analysis/requests/fe/pdf/lng.pdf*.

Eydeland, A. and Wolyniec, K., 2003. Energy and power risk management: New developments in modeling, pricing, and hedging (Vol. 206). John Wiley & Sons.

Fisher, F.M. and Kaysen, C., 1962. A Study in Econometrics: The Consumption for Electricity in the US Amsterdam.

Geman, H. and Ohana, S., 2009. Forward curves, scarcity and price volatility in oil and natural gas markets. *Energy Economics*, *31*(4), pp.576-585.

Gouel, C. and Legrand, N., 2017. Estimating the competitive storage model with trending commodity prices. *Journal of Applied Econometrics*, *32*(4), pp.744-763.

Hausman, C. and Kellogg, R., 2015. Welfare and distributional implications of shale gas (No. w21115). *National Bureau of Economic Research*.

Hochradl, M. and Rammerstorfer, M., 2012. The convenience yield implied in European natural gas hub trading. *Journal of Futures Markets*, 32(5), pp.459-479.

Höffler, F. and Kübler, M., 2007. Consumption for storage of natural gas in northwestern Europe: Trends 2005–30. *Energy Policy*, 35(10), pp.5206-5219.

Houthakker, H.S., Verleger, P.K. and Sheehan, D.P., 1974. Dynamic consumption analyses for gasoline and residential electricity. *American Journal of Agricultural Economics*, *56*(2), pp.412-418.

Hulshof, D., van der Maat, J.P. and Mulder, M., 2016. Market fundamentals, competition and natural-gas prices. *Energy Policy*, *94*, pp.480-491.

Judd, K.L., 1998. *Numerical methods in economics*. MIT press.

Kaldor, N., 1939. Speculation and economic stability. *The Review of Economic Studies*, *7*(1), pp.1-27.

Kanamura, T., 2009. A supply and consumption based volatility model for energy prices. *Energy Economics*, *31*(5), pp.736-747.

Karl, T. and Koss, W.J., 1984. Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983. National Climatic Data Center.

Kong, Z., Dong, X. and Zhou, Z., 2015. Seasonal imbalances in natural gas imports in major northeast Asian countries: variations, reasons, outlooks and countermeasures. *Sustainability*, *7*(2), pp.1690-1711.

Kremser, T. and Rammerstorfer, M., 2010, October. Convenience yield and risk premium—comparison of the European and US natural gas markets. *In Proceedings of the 23rd Australasian Finance and Banking Conference* (Vol. 11, pp. 1-42).

Lence, S.H. and Hayes, D.J., 2002. US farm policy and the volatility of commodity prices and farm revenues. *American Journal of Agricultural Economics*, 84(2), pp.335-351.

Levi, M.A., 2012. *A Strategy for US Natural Gas Exports*. Hamilton Project, Brookings Institution.

Lin, C.Y.C. and Prince, L., 2013. Gasoline price volatility and the elasticity of consumption for gasoline. *Energy Economics*, *38*, pp.111-117.

Makki, S.S., Tweeten, L.G. and Miranda, M.J., 1996. Wheat storage and trade in an efficient global market. *American Journal of Agricultural Economics*, *78*(4), pp.879-890.

Makki, S.S., Tweeten, L.G. and Miranda, M.J., 2001. Storage–trade interactions under uncertainty: Implications for food security. *Journal of Policy modeling*, *23*(2), pp.127-140.

Medlock III, K.B., 2012. US LNG exports: truth and consequence. Institute For public Policy Research. Rice University

Miranda, M.J. and Glauber, J.W., 1993. Estimation of dynamic nonlinear rational expectations models of primary commodity markets with private and government stockholding. *The Review of Economics and Statistics*, pp.463-470.

Miranda, M.J., 1997. Numerical strategies for solving the nonlinear rational expectations commodity market model. *Computational Economics*, 11(1-2), pp.71-87.

Miranda, M.J. and Fackler, P.L., 2004. *Applied computational economics and finance*. MIT press.

Miranda, M.J. and Fackler, P.L., 2007. Computational Methods in Economics-MATLAB Toolbox. File downloadable from the website *http://www4.ncsu.edu/unity/users/p/pfackler/www/ECG790C*.

Miranda, M.J. and Rui, X., 1999. An empirical reassessment of the commodity storage model. *Unpublished working paper, Ohio State University*.

MIT, 2010. The future of natural gas: an interdisciplinary MIT study interim report. *Massachusetts Institute of Technology, Cambridge, MA.*

Mohammadi, H., 2011. Market integration and price transmission in the US natural gas market: From the wellhead to end use markets. *Energy Economics*, 33(2), pp.227-235.

Moryadee, S., Gabriel, S.A. and Avetisyan, H.G., 2014. Investigating the potential effects of US LNG exports on global natural gas markets. *Energy Strategy Reviews*, *2*(3), pp.273-288.

Mu, X., 2007. Weather, storage, and natural gas price dynamics: Fundamentals and volatility. *Energy Economics*, 29(1), pp.46-63.

NERA Economic Consulting, 2012. Macroeconomic impacts of LNG exports from the United States. Available from: *https://energy.gov/sites/prod/files/2013/04/f0/nera_lng_report.pdf*

Nick, S. and Thoenes, S., 2014. What drives natural gas prices?—A structural VAR approach. *Energy Economics*, *45*, pp.517-527.

Paltsev, S., Jacoby, H.D., Reilly, J.M., Ejaz, Q.J., Morris, J., O'Sullivan, F., Rausch, S., Winchester, N. and Kragha, O., 2011. The future of US natural gas production, use, and trade. *Energy Policy*, *39*(9), pp.5309-5321.

Peterson, H.H. and Tomek, W.G., 2005. How much of commodity price behavior can a rational expectations storage model explain?. *Agricultural Economics*, 33(3), pp.289-303.

Pindyck, R.S., 2001. The dynamics of commodity spot and futures markets: a primer. *The energy journal*, pp.1-29.

Pindyck, R.S., 2004. Volatility in natural gas and oil markets. *Journal of Energy and Development* 30 (1), 1–19.

Pirrong, C., 2011. *Commodity price dynamics: A structural approach*. Cambridge University Press.

Ratner, M., Parfomak, P.W., Fergusson, I.F. and Luther, L., 2013. *US Natural Gas Exports: New Opportunities, Uncertain Outcomes*. Congressional Research Service.

Ritz, R.A., 2014. Price discrimination and limits to arbitrage: An analysis of global LNG markets. *Energy Economics*, *45*, pp.324-332.

Roberts, M.J. and Schlenkera, W., 2013. Identifying supply and consumption elasticities of agricultural commodities: Implications for the US ethanol mandate. *The American Economic Review*, 103(6), pp.2265-2295.

Rui, X. and Miranda, M.J., 1995. Commodity storage and interseasonal price dynamics. In *NCR-134 Conference Proceedings* (pp. 44-53).

Sánchez-Úbeda, E.F. and Berzosa, A., 2007. Modeling and forecasting industrial end-use natural gas consumption. *Energy Economics*, 29(4), pp.710-742.

Schoppe, J., 2010. The valuation of natural gas storage: a knowledge gradient approach with non-parametric estimation (Doctoral dissertation, Princeton University).

Thompson, M., Davison, M. and Rasmussen, H., 2009. Natural gas storage valuation and optimization: A real options application. *Naval Research Logistics (NRL)*, 56(3), pp.226-238.

Tran, A.N., 2013. Essays on Price Adjustment in a Competitive Storage Model. North Carolina State University.

U.S. Department of Energy, 2015. The macroeconomic impact of increasing U.S. LNG exports. Available from: *https://energy.gov/sites/prod/files/2015/12/f27/20151113_macro_impact_of_lng_exports_0.pdf*.

van Goor, H. and Scholtens, B., 2014. Modeling natural gas price volatility: The case of the UK gas market. *Energy*, *72*, pp.126-134.

Vivoda, V., 2014. LNG import diversification in Asia. *Energy Strategy Reviews*, *2*(3), pp.289-297.

Volmer, T., 2011. A robust model of the convenience yield in the natural gas market. *Journal of Futures Markets*, 31(11), pp.1011-1051.

Wei, S.Z.C. and Zhu, Z., 2006. Commodity convenience yield and risk premium determination: The case of the US natural gas market. *Energy Economics*, *28*(4), pp.523-534.

Wiggins, S. and Etienne, X.L., 2017. Turbulent times: Uncovering the origins of US natural gas price fluctuations since deregulation. *Energy Economics*, *64*, pp.196-205.

Williams, J.C. and Wright, B.D., 1991. Storage and commodity markets. Cambridge university press.

Working, H., 1949. The theory of price of storage. *The American Economic Review*, *39*(6), pp.1254-1262.

APPENDIX

Table A1a. Residential Sector Consumption Estimation - First Stage Result*

| Winter ( Dec. - Feb.) | | | Summer ( May. - Sep.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| IV-Lagged Weather | 0.01 | 0.00 | IV-Lagged Weather | 0.004 | 0.001 | IV-Lagged Weather | 0.004 | 0.001 |
| log(D_p) | 0.01 | 0.13 | log(D_p) | 0.30 | 0.09 | log(D_p) | 0.36 | 0.09 |
| log (HDD) | 1.01 | 0.14 | | | | log(HDD) | 0.20 | 0.06 |
| Adj. R-squared | 0.19 | | Adj. R-squared | 0.12 | | Adj. R-squared | 0.14 | |
| F-Statistic | 9.76 | | F-Statistic | 9.09 | | F-Statistic | 8.98 | |

Table A1b. Commercial Sector Consumption Estimation - First Stage Result*

| Winter ( Dec. - Feb.) | | | Summer ( May. - Sep.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| IV-Lagged Weather | 0.01 | 0.00 | IV-Lagged Weather | 0.005 | 0.001 | IV-Lagged Weather | 0.005 | 0.001 |
| log(D_p) | -0.48 | 0.16 | log(D_p) | -0.05 | 0.10 | log(D_p) | 0.12 | 0.09 |
| log(HDD) | 1.08 | 0.14 | | | | log(HDD) | 0.18 | 0.06 |
| Adj. R-squared | 0.20 | | Adj. R-squared | 0.11 | | Adj. R-squared | 0.12 | |
| F-Statistic | 10.57 | | F-Statistic | 8.23 | | F-Statistic | 7.79 | |

Table A1c. Industrial Sector Consumption Estimation - First Stage Result*

| Winter (Nov. - Mar.) | | | Summer ( Jun. - Aug.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| IV-Lagged Weather | 0.01 | 0.00 | IV-Lagged Weather | 0.01 | 0.00 | IV-Lagged Weather | 0.01 | 0.00 |
| log(D_p) | -1.19 | 0.18 | log(D_p) | -1.03 | 0.24 | log(D_p) | -0.62 | 0.20 |
| log(HDD) | 0.44 | 0.10 | | | | log(HDD) | 0.03 | 0.04 |
| Adj. R-squared | 0.25 | | Adj. R-squared | 0.22 | | Adj. R-squared | 0.22 | |
| F-Statistic | 12.07 | | F-Statistic | 7.78 | | F-Statistic | 8.76 | |

Table A1d. Electrical Sector Consumption Estimation - First Stage Result*

| Winter ( Nov. - Mar.) | | | Summer ( Jun. - Aug.) | | | Shoulder Season | | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | Std. Error | | Estimate | Std. Error | | Estimate | Std. Error |
| IV-Lagged Weather | 0.005 | 0.001 | IV-Lagged Weather | 0.005 | 0.001 | IV-Lagged Weather | 0.01 | 0.00 |
| log(D_p) | -0.19 | 0.05 | log(D_p) | -0.20 | 0.06 | log(D_p) | -0.15 | 0.05 |
| log(HDD) | 0.54 | 0.09 | log(CDD) | 0.06 | 0.06 | log(CDD) | -0.03 | 0.03 |
| Adj. R-squared | 0.32 | | Adj. R-squared | 0.21 | | Adj. R-squared | 0.23 | |
| F-Statistic | 13.52 | | F-Statistic | 6.07 | | F-Statistic | 7.98 | |

*Note: 1) Table A1a to table A1d above report the first stage estimates for the 2SLS estimates presented in Chapter 3, table 4 to table 7;
2) The independent variable for regression in Table A1a to table A1c is natural gas real price in log. The independent variable in Table A1d is the natural gas to coal price ratio in log;
3) IV- Lagged Weather denotes the instrument variable used for consumption estimation. The Instrument Variable utilized in this study is the accumulated lagged weather variables in other regions; log(D_p) represents the consumption level in prior month in log; HDD and CDD are the current month heating degree days and cooling degree days;
4) The second stage regression excludes instrument variable.